PHD

Branching diffusions

Hardy, Robert

Award date:
2003

Awarding institution:
University of Bath

Branching Diffusions

submitted by

Robert Hardy

for the degree of Ph.D.

of the

University of Bath

2003

All the work that has gone into this Ph.D has only been possible due to the continued support and love of Rossana, and I dedicate this thesis to her.

Summary

Within the large body of theories on branching Brownian motion (BBM) and typed branching diffusions this thesis presents work that focuses mainly on families of strictly-positive additive martingales and changes of measure, and further develops and uses the recent spine approaches to obtain results on the $\mathcal{L}^p$-convergence of these martingales as well as strong results in the large-deviations theory of branching diffusions. The thesis is divided into three related parts: in the first part we present earlier work on a classical approach to branching diffusions, where a single-particle model inspired by the work of Harris and Williams [21] together with Perron-Frobenius theory are used to consider the conditions for $\mathcal{L}^p$-convergence of an additive martingale for a finitely-typed branching diffusion. A main theme of this thesis is to show that spine-based techniques can give much better proofs for these questions, and in the second part of this thesis, after laying out a new formulation that substantially improves the existing spine approach, we use spines to consider the $\mathcal{L}^p$-convergence of additive martingales in three different models of branching diffusions. In the third and final part of this thesis we apply spines to the theory of large-deviations for branching diffusions: in Chapter 5 we use our new formulation of the spine foundations to introduce a new and very general class of additive martingales for BBM and prove a large-deviations principle for BBM that is analogous to Schilder's theorem for Brownian motion, and in Chapter 6 we prove an important lower bound for large deviations in a typed branching diffusion originally studied by Harris and Williams [21], and here the proof likewise centres on martingale estimates via the spine decomposition. The techniques that we develop here should have the added benefit of being applicable to martingales and large-deviations problems for a much more general class of branching-diffusion models.

Acknowledgements

I would like to thank my supervisor Dr Simon Harris for his extremely valuable contributions and intuition throughout the course of the Ph.D, and especially during those difficult times when progress seemed nearly impossible. I wish him all the best for the future.

Thanks also to all those people in the Department of Mathematical Sciences at the University of Bath for their everyday help, and in particular the secretarial staff, my officemates and computing support. Particular best wishes to my probability colleagues John, Pascal, Alex and Ramses, Peter, Chris and David.

Finally, I am grateful to my funding council *EPSRC* for the financial support that made it possible for me to carry out this work.

Contents

1 Outline of Thesis **6**

I A classical approach to $\mathcal{L}^p$-convergence 10

2 A typed branching diffusion **11**

2.1 The model description 11
2.2 Martingales for the branching diffusion 13
2.3 The *Single-Particle Model* and *Many-to-One* theorem 15
2.3.1 A single-particle martingale and measure change 16
2.4 Perron-Frobenius Theory for $\frac{1}{2}\lambda^2 A + \theta Q + R$ and E_λ 21
2.5 A first proof of $\mathcal{L}^p$ convergence 27
2.5.1 Part 1: $\lambda \in (\tilde{\lambda}(\theta), 0]$ 27
2.5.2 Part 2: $\lambda < \tilde{\lambda}(\theta)$ 29
2.6 Travelling waves and FKPP problems 31
2.6.1 The FKPP equation 31
2.6.2 Relation to BBM and the McKean Representation 32
2.6.3 FKPP systems for the finite-type model 34
2.6.4 Martingales and Travelling Waves 34
2.6.5 Results on Monotone Travelling Waves 35

II A new spine approach to $\mathcal{L}^p$ convergence 44

3 A new formulation of the spine approach **45**

3.1 The underlying space for spines 49
3.1.1 Marked Galton-Watson trees with spines 49
3.1.2 Filtrations 52
3.2 Probability measures 53
3.3 Martingales 55
3.4 Changing the measures 59
3.4.1 Understanding the measure $\tilde{\mathbb{Q}}$ 60

3.4.2 A discussion of the martingales and new measures 62
3.5 The spine decomposition 62
3.6 New spine results 65
3.6.1 The *Gibbs-Boltzmann* weights of $\tilde{\mathbb{Q}}$ 65
3.6.2 Conditional expectations and the Kesten-Stigum Theorem 67
3.6.3 The *Full* Many-to-One Theorem 68

4 Spine proofs for $\mathcal{L}^p$-convergence **70**
4.1 Overview 70
4.2 Branching Brownian motion 71
4.2.1 The $\mathcal{L}^1$-convergence theorem of Kyprianou 72
4.2.2 A spine proof of $\mathcal{L}^p$-convergence 75
4.2.3 Random family sizes 79
4.3 The finite-type branching diffusion 81
4.3.1 The spine process (ξ_t, η_t) under $\tilde{\mathbb{Q}}_\lambda$ 82
4.3.2 Proof of Theorem 4.3.1 84
4.3.3 Random family sizes 86
4.4 A continuous-type branching diffusion 89
4.4.1 The measure change 89
4.4.2 Proof of Theorem 4.4.2 91

III Large-deviations via spines 93

5 A large-deviations result for BBM **94**
5.1 Introduction and outline of proof 94
5.2 A local upper bound 97
5.3 A new martingale Z_{g_T} for BBM 99
5.4 The growth of Z_{g_T} under $\mathbb{Q}_{g_T}$ 101
5.5 A local lower bound 103
5.6 Improving the 'weak' large-deviations result 104
5.7 A brief discussion of $S(g)$ 108

6 Large deviations for the HW model **110**
6.1 Introduction and statement of result 110
6.2 Review of the Harris and Git almost-sure result 111
6.2.1 The almost-sure upper bound 112
6.2.2 A two-phase mechanism for the lower bound 113
6.3 Large-deviations heuristics 114
6.4 Martingales and measures 117
6.5 Proving the large-deviations lower bound 121
6.6 A spine proof of the martingale upper-bound 122

Chapter 1

Outline of Thesis

Within the large body of theories on branching Brownian motion (BBM) and typed branching diffusions this thesis presents work that focuses mainly on families of strictly-positive additive martingales and changes of measure, and further develops and uses the recent spine approaches to obtain results on the $\mathcal{L}^p$-convergence of these martingales as well as strong results in the large-deviations theory of branching diffusions. The thesis is divided into three related parts: in the first part we present earlier work on a classical approach to branching diffusions, where a single-particle model inspired by the work of Harris and Williams [21] together with Perron-Frobenius theory are used to consider the conditions for $\mathcal{L}^p$-convergence of an additive martingale for a finitely-typed branching diffusion. A main theme of this thesis is to show that spine-based techniques can give much better proofs for these questions, and in the second part of this thesis, after laying out a new formulation that substantially improves the existing spine approach, we use spines to consider the $\mathcal{L}^p$-convergence of additive martingales in three different models of branching diffusions. In the third and final part of this thesis we apply spines to the theory of large-deviations for branching diffusions: in Chapter 5 we use our new formulation of the spine foundations to introduce a new and very general class of additive martingales for BBM and prove a large-deviations principle for BBM that is analogous to Schilder's theorem for Brownian motion, and in Chapter 6 we prove an important lower bound for large deviations in a typed branching diffusion originally studied by Harris and Williams [21], and here the proof likewise centres on martingale estimates via the spine decomposition. The techniques that we develop here should have the added benefit of being applicable to martingales and large-deviations problems for a much more general class of branching-diffusion models.

Work in this thesis, joint with Dr S.C.Harris, appears also in articles [16], [17], [18], [19] and [20].

The martingales and spine techniques

In most branching models a strictly-positive martingale can be defined via a sum over the particles alive at time t, and in general this relates closely to obtaining a strictly-positive solution of an associated eigenvalue problem. In different contexts these martingales have been considerably studied, and we note the work of Biggins [2] on $\mathcal{L}^1$ convergence of the martingale in branching random walks, Asmussen and Hering [1] on $\mathcal{L}^1$ and $\mathcal{L}^p$ $(p > 1)$ convergence for general typed branching processes, Neveu [44] on $\mathcal{L}^p$ convergence for BBM, and more recently Harris and Williams [21] on $\mathcal{L}^p$ convergence for an interesting example of a typed branching diffusion.

All these above references can be said to use a *classical* approach, by which it is meant that the standard branching decomposition and the expectation semigroup form the main axis of the proofs, and in section 2.5 we present a classical proof (which generalizes the work Champneys *et al* [4] and relates very closely to the model and proofs of Harris and Williams [21]) of the $\mathcal{L}^p$ convergence of the martingale in a specific model of typed branching diffusion where each particle's type is governed by a finite-state, time-reversible Markov chain.

This opening study gives us the opportunity in section 2.3 to discuss at a more abstract level some of the important techniques used by Harris and Williams [21] such as the *single-particle model* and their use of a change-of-measure. For this finite-state model, the related eigenvalue problem can be tackled with Perron-Frobenius theory and we dedicate the whole of section 2.4 to giving new proofs regarding the behaviour of the largest eigenvalue E_λ as a function of a model parameter $\lambda \in \mathbb{R}$. Specific properties of E_λ such as its convexity or its derivative turn out to be crucially important in determining the convergence behaviour of the martingale, as was already seen in the Champneys *et al* model, and the Harris and Williams model. As an adjunct to this section, giving some context to the main themes of the thesis, we also include a discussion of the travelling waves that are typically related to branching diffusions, and in the case of this finite-type model we include new proofs of an asymptotic result first proven in an analogous form by Harris [23] for BBM using *probabilistic* techniques only.

Spine techniques have already given significantly clearer proofs of $\mathcal{L}^1$ convergence properties of the martingales in different models of branching processes. Examples are the proofs of $\mathcal{L}^1$ convergence for the BBM martingale by Kyprianou [35], and the analogous case for Galton-Watson processes proved by Lyons [40]. Other spine-based work includes Kyprianou and Sani [36], Biggins and Kyprianou [3].

Spines have not previously been used for proofs of $\mathcal{L}^p$ convergence and in chapter 4 this is the main focus – we give spine proofs of the $\mathcal{L}^p$ convergence of martingales for three different models of branching diffusions (in two of these we also give proofs for a variation of each model that allows *random* family sizes). There are a number of reasons why we may be interested in knowing about the $\mathcal{L}^p$ convergence of a martingale: in Neveu's original article [44] it was a means to proving $\mathcal{L}^1$-convergence, whilst Harris and Git [25] and Asmussen and Hering [1] have used it to deduce the almost-sure rate of convergence of the martingale to its limit. Of equal importance are the *techniques* that we use here, and similar ideas will be developed in the final part of this thesis in the context of the large-deviations theory of branching diffusions.

However, before embarking on these proofs for $\mathcal{L}^p$ convergence, in Chapter 3 we present our improved version of the spine approach which not only successfully repairs a weak point in the approaches based on the Lyons *et al* papers [40, 34, 41] – namely, we show how to build three new measures which are all *probability measures*, whereas the Lyons *et al* configuration produced only one – but which also brings out many more relationships between the three change-of-measure martingales, since we use different *filtrations* rather than different underlying spaces as has previously been the case.

This new formulation gives substantial improvements to both the Harris and Williams single-particle approach and the previous spine approaches (exemplified in Lyons *et al* [40, 34, 41]), since it combines them into a a single and more powerful object which can be interpreted from either point of view. In the first instance this gives us the correct formal basis in which to express the relationship between the single-particle martingale ζ_λ and the additive martingale Z_λ that we used in Chapter 2. This same idea will work in more generality to give us a consistent methodology for developing *new martingales* for branching diffusions that can be very powerful in proving large-deviations results, as we explore fully in Part III.

A new and interesting aspect of our formulation is the relation that becomes clear between the spine and the 'Gibbs-Boltzmann' weightings for the branching particles, which can be seen as conditional expectations of spine events. This is further developed to show how such conditional expectations can obtain a new and very useful interpretation of the additive operations previously seen within the context of the Kesten-Stigum theorem and related problems. Furthermore, as a consequence of the Gibbs-Boltzmann weightings we obtain a substantially easier proof of an improved version of the Many-to-One theorem used by Harris and Williams [21].

Large deviations results via spines

In the third part of the thesis we present another new application of spine techniques in the area of large deviations for branching diffusions.

Large-deviations principles for a single diffusing particle have been known for a long time, Schilder's theorem being the example for Brownian motion. As far as *branching* Brownian motion is concerned, the natural generalization of Schilder's theorem would be to give an exponential rate of the decay of the probability that *at least one* particle follows a certain path. In fact, for an upper-bound of this probability a simple estimate against an expectation calculation gives the result (this uses our improved Many-to-One theorem), but in general the lower-bound is more difficult since classical techniques cannot easily force one of the BBM particles to follow a chosen path, which is a common change-of-measure technique in large deviations.

Tzong-Yow Lee [37] proved such a large-deviations principle for BBM but for the lower bound relied heavily on Freidlin's work on rescalings of solutions of reaction-diffusion equations – it could actually be said that most of the work was already contained in Friedlin's results. In contrast to this, the spine approach is very natural since we immediately have the change of measure required to force the *spine* itself along a path, and we give a very clear spine-based proof of the full large-deviations principle for BBM in chapter 5. An important part of our proof consists in obtaining an upper bound for the growth of the martingale under the new measure (where it becomes a *sub*martingale), and we do this using Doob's submartingale inequality along with spine techniques developed from those used in the chapter on $\mathcal{L}^p$-convergence.

In Harris and Git [25] a result is stated for a model of typed-branching diffusion regarding the almost-sure number of particles at large space-type locations, which required a difficult large-deviation lower bound on the probability that at least one of the diffusing particles in the branching process will be near a specific large space-type location at some fixed time, given that the original ancestor started at the space-type origin. In Chapter 6 we give a spine proof of the difficult lower-bound for a large-deviations event – the martingale ideas that we used for the branching Brownian motion case are here substantially improved and combined with applications of Varadhan's lemma to obtain an upper bound on the exponential growth rate of a new martingale for this model. Through a change of measure with this martingale we obtain the lower bound for the large deviations, and briefly discuss how we could expect to also obtain an upper bound using spine ideas in terms of the Many-to-One theorem. Again, the spine techniques that we develop have the benefit of offering a consistent approach to large-deviation problems and it should be possible to apply them to more general models of branching diffusions. Furthermore, through the *exact* spine decomposition of the martingale they give a methodology for avoiding the difficult non-linear calculations that were necessary in the estimates used for the classical approaches such as Harris and Williams [21]

Part I

A classical approach to $\mathcal{L}^p$-convergence

Chapter 2

A typed branching diffusion

In Champneys *et al* [4] a 2-type model of branching diffusion was considered, and they gave proofs regarding the $\mathcal{L}^p$-convergence of an additive martingale defined in terms of the branching diffusion. In this chapter we consider a generalization of their model and use techniques that have also been effectively used by Harris and Williams [21] to deal with the issue of $\mathcal{L}^p$-convergence of the martingale. Briefly, these 'classical' approaches (in contrast to our spine approaches in later chapters) are based mainly on a useful inequality stated in Neveu [44] which combines with a 'Many-to-One' idea of converting expectation calculations over the whole collection of branching particles to an appropriately up-weighted similar expectation involving only a single 'typical' particle. To carry out these single-particle expectation calculations, we then use the idea of Harris and Williams [21] to change the measure via a single-particle martingale closely related to the additive martingale for the whole collection of branching particles.

In section 2.6 we present some work on the relationship between branching diffusions and a class of partial-differential equations known as FKPP equations. This body of work, carried out using the classical approaches, is offered as an adjunct to the main themes of the thesis to give context and show some new results that were obtained as part of earlier work on the finite-type model.

2.1 The model description

Suppose that for a fixed $n \in \mathbb{N}$ we are given two sets of positive constants $a(1), \ldots, a(n)$ and $r(1), \ldots, r(n)$. Consider a typed branching diffusion in which the type of each particle moves as a finite, irreducible and time-reversible Markov chain on the set $I := \{1, \ldots, n\}$ with Q-matrix θQ (θ is a strictly positive constant that could be considered as the *temperature* of the system) and invariant measure $\pi = (\pi(1), \ldots, \pi(n))$. The spatial movement of a particle of type y is a driftless Brownian motion with instantaneous variance $a(y)$, so that if $(X_u(t), Y_u(t)) \in \mathbb{R} \times I$ is the space-type location of individual u at time t then we have

$$\mathrm{d}X_u(t) = a(Y_u(t))\,\mathrm{d}B_t$$

for a Brownian motion B_t. Fission of a particle of type y occurs at a rate $r(y)$ to produce two particles at the same space-type location as the parent.

We define $J := \mathbb{R} \times I$, and suppose that the configuration of this whole branching diffusion at time t is given by the J-valued point process $\mathbb{X}_t = \{(X_u(t), Y_u(t)) : u \in N_t\}$ where N_t is the set of individuals alive at time t. Let the measures $\{P^{x,y} : (x,y) \in \mathbb{R}^2\}$ on the filtered space $(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0})$ be such that under $P^{x,y}$ the initial ancestor starts at (x,y) and $\mathbb{X}_t$ becomes the above-described branching diffusion – that is, we are supposing that $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration generated by the point process $\mathbb{X}_t$.

Throughout this thesis we use the *de Finetti* notation in which a measure symbol P will stand both for the probability measure and the expectation operator; see [49] for a clear discussion of this.

This branching diffusion is a generalization of a 2-type model studied in Champneys *et al* [4], which itself was inspired by a similar but more complex model laid out by Harris and Williams [21] in which the type process moves as an Orstein-Uhlenbeck process on $\mathbb{R}$ with the diffusion and fission coefficients $a(y)$ and $r(y)$ being quadratic functions of the type y. This model will be considered in section 4.4 and later in chapter 6.

It should be noted that the condition of time-reversibility on the Markov chain is not absolutely necessary, and is really just a simplifying assumption that gives us an easier $\mathcal{L}^2$ theory for the matrices and eigenvectors. If we were to drop this assumption we would still have an $\mathcal{L}^2$ theory since we are working with finite vectors, but it would not be so immediate.

Definition 2.1.1 *For two vectors u, v on I, we use the invariant measure π to define an inner product:*

$$\langle u, v\rangle_\pi := \sum_{i=1}^{n} u_i v_i \pi_i,$$

which gives us a Hilbert space of vectors on I which we refer to as $\mathcal{L}^2(\pi)$.

An alternative way of writing this inner product is:

$$\langle u, w\rangle_\pi = u^T \Pi w,$$

where Π is the diagonal matrix $\mathrm{diag}[\pi_1, \ldots, \pi_n]$. The time-reversibility assumption is entirely motivated by the following simple result

Theorem 2.1.2 *The matrix θQ is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\pi$.*

Proof: Because $\langle Qg, h\rangle_\pi = g^T Q^T \Pi h = g^T (\Pi Q)^T h$, the self-adjoint property holds if and only if $(\Pi Q)^T = \Pi Q$, and this means:

$$\pi_i q_{ij} = (\Pi Q)_{ij} = (\Pi Q)_{ji} = \pi_j q_{ji},$$

which is exactly the condition of time-reversibility. □

We gather together the constants as $n \times n$ matrices:

$$A := \mathrm{diag}[a(1), \ldots, a(n)], \quad R := \mathrm{diag}[r(1), \ldots, r(n)],$$

and because they are diagonal we can immediately deduce the following corollary:

Corollary 2.1.3 *The matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$ is self-adjoint with respect to $\langle\cdot,\cdot\rangle_\pi$.*

2.2 Martingales for the branching diffusion

Via generators we show below that for any $\lambda \in \mathbb{R}$, any function (vector) $v_\lambda : I \to \mathbb{R}$ and any number $E_\lambda \in \mathbb{R}$, the expression

$$Z_\lambda(t) := \sum_{u \in N_t} v_\lambda(Y_u(t))\, e^{\lambda X_u(t) - E_\lambda t}, \tag{2.1}$$

will be a local martingale if v_λ and E_λ satisfy:

$$\Big(\frac{1}{2}\lambda^2 A + \theta Q + R\Big) v_\lambda = E_\lambda v_\lambda, \tag{2.2}$$

which is to say that v_λ must be an eigenvector of the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$, with eigenvalue E_λ. We suppose that the eigenvector v_λ is normalized so that $\|v_\lambda\|_\pi := \langle v_\lambda, v_\lambda\rangle_\pi = 1$.

The fact that the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$ is self-adjoint is enough to guarantee the existence of eigenvectors in $\mathcal{L}^2(\pi)$, but the fact that we are dealing with a finite-state Markov chain means that we also have the *Perron-Frobenius* theory to hand, which allows us to suppose that v_λ is a *strictly positive* eigenvector whose eigenvalue E_λ is real and the farthest to the right of all the other eigenvalues. Perron-Frobenius theory is discussed in detail in section 2.4.

After we have developed some more theory, we shall in section 2.3 show that $Z_\lambda(t)$ is actually a *true* martingale when condition (2.2) holds, not just local. For now we state this as a theorem but here prove only the *local* martingale property.

Theorem 2.2.1 *For each $\lambda \in \mathbb{R}$,*

$$Z_\lambda(t) = \sum_{u \in N_t} v_\lambda(Y_u(t))\, e^{\lambda X_u(t) - E_\lambda t}, \tag{2.3}$$

defines a strictly positive (and hence a.s. convergent) martingale, where E_λ is the (real) largest eigenvalue and v_λ is the corresponding normalized, strictly positive eigenvector of the self-adjoint $n \times n$ matrix

$$\frac{1}{2}\lambda^2 A + \theta Q + R.$$

Proof: The state space of the branching diffusion can be identified with the set

$$S := \bigcup_{n \geq 1} \left(\{n\} \times \mathbb{R}^n \times I^n\right),$$

so that the state of the process at a time t is described in terms of the array

$$\big(|N_t|, \mathbf{X}(t), \mathbf{Y}(t)\big) := \big(|N_t|; (X_u(t) : u \in N_t); (Y_u(t) : u \in N_t)\big)$$

On this space the formal generator $\mathcal{G}$ for the branching diffusion can be written as:

$$\mathcal{G} = \mathcal{G}_A + \mathcal{G}_Q + \mathcal{G}_R,$$

where for $n \geq 1$, $x \in \mathbb{R}^n$ and $y \in I^n$ we have (for $F : S \to \mathbb{R}$)

$$\begin{aligned}
\mathcal{G}_A F(n,x,y) &= \sum_{k=1}^{n} \frac{1}{2} a(y_k) \frac{\partial^2 F}{\partial x_k^2}, \\
\mathcal{G}_Q F(n,x,y) &= \theta \sum_{k=1}^{n} \sum_{j \neq y_k} Q(y_k, j) \{F(n,x,s_{k,j}(y)) - F(n,x,y)\}, \\
\mathcal{G}_R F(n,x,y) &= \sum_{k=1}^{n} r(y_k) \{F(n+1,(x,x_k),(y,y_k)) - F(n,x,y)\},
\end{aligned}$$

where $s_{k,j}(y) := (y_1, \ldots, y_{k-1}, j, y_{k+1}, \ldots, y_n)$ and $(x, x_k) := (x_1, \ldots, x_n, x_k)$, etc.

The following well-known result can be found in Rogers and Williams [51]):

Lemma 2.2.2 *If $F : [0,\infty) \times S \to \mathbb{R}$, then $F(t, N_t, \mathbf{X}(t), \mathbf{Y}(t))$ is a* local martingale *if and only if F is in the kernel of the operator (space-time generator) $\partial/\partial t + \mathcal{G}$:*

$$\Big(\frac{\partial}{\partial t} + \mathcal{G}\Big) F(t,n,x,y) = 0 \qquad (n \geq 1, x \in \mathbb{R}^n, y \in I^n). \tag{2.4}$$

An application of this lemma implies that the additive expression $Z_\lambda(t)$ defined at (2.1) will be a local martingale if the relation (2.2) is satisfied. Perron-Frobenius theory guarantees the existence of v_λ and E_λ, and therefore we have shown that Z_λ is a strictly-positive local martingale (which in fact immediately implies that it is a supermartingale). **Remark:** As mentioned, the proof that it is a true martingale will be completed at Corollary 2.3.4.

□

Taking the martingale property as stated, we can be sure that the limit

$$Z_\lambda(\infty) := \lim_{t \to \infty} Z_\lambda(t)$$

exists and is finite $P^{x,y}$-almost surely. It is typical to have a '0-1' situation for the limit variable $Z_\lambda(\infty)$; the following proof was stated by Harris and Williams [21] but will apply to many models including our own:

Theorem 2.2.3 *For all $\lambda \in \mathbb{R}$,*

$$P^{x,y}(Z_\lambda(\infty) = 0) = 0 \text{ or } 1.$$

Proof: First we note that for $\lambda \in \mathbb{R}$,

$$P^{x,y}\big(Z_\lambda(\infty) = 0\big) = P^{0,y}(e^{\lambda x} Z_\lambda(\infty) = 0) = P^{0,y}(Z_\lambda(\infty) = 0) =: g(y).$$

Thus it is clear that the probability is independent of the spatial starting position. Moreover, for all $t \geq 0$, $y \in \mathbb{R}$,

$$\begin{aligned}
g(y) &= P^{0,y}(Z_\lambda(\infty) = 0) = P^{0,y}\left(P\big(Z_\lambda(\infty) = 0 | \mathcal{F}_t\big)\right) \\
&= P^{0,y}\left(\prod_{k=1}^{|N(t)|} g(Y_k(t))\right) \leq P^{0,y}\left(g(Y_1(t))\right).
\end{aligned}$$

Hence $g(Y_1(t))$ is a bounded submartingale which therefore converges. Since Y_1 is recurrent, this convergence implies that $g(\cdot)$ must be constant on the state space I and hence, because of the equality with a product above, $g(\cdot) = 0$ or 1:

$$\text{for } any\ \lambda \in \mathbb{R}, \qquad P^{x,y}(Z_\lambda(\infty) = 0) \in \{0, 1\}.$$

$\square$

Thus the martingale $Z_\lambda(t)$ either converges to a strictly-positive limit $Z_\lambda(\infty)$, or it converges to 0.

In the case that the limit $Z_\lambda(\infty)$ is strictly positive, the question of which further conditions on the model might be required to ensure that this convergence also occurs in $\mathcal{L}^p$ for some $p > 1$ has been addressed in a number of different branching models. We shall see that for our model the condition

$$pE_\lambda - E_{p\lambda} > 0, \qquad \text{for some } p \in (1, 2], \tag{2.5}$$

will be sufficient to guarantee that the martingale is $\mathcal{L}^p$-convergent. In fact, for this finite-type model we show in section 2.4 that the eigenvalue E_λ is a strictly-convex function of λ, and this translates to mean that there is a bounded, open interval about 0 such that for each λ within this interval we can be sure that there is always some $p \in (1, 2]$ (and therefore many) such that condition (2.5) holds. Thus we deduce that for all λ inside an open, bounded interval about 0 the martingales $Z_\lambda(t)$ are convergent in $\mathcal{L}^1$ to a strictly-positive limit. On the other hand we shall be able to show that for all λ outside this interval the martingale limit is null; we note that our proofs will not cover the boundary points where different techniques are needed based on 'derivative' martingales (see Kyprianou [35] and Harris [23] for examples). This is a typical feature of a branching diffusion and can most easily be seen as relating to the fact that there is an upper bound to the velocity at which a branching diffusion spreads out over $\mathbb{R}$ in the long term, or equivalently that *travelling-waves* defined from the martingales have a minimum possible speed; see section 2.6 for details of the relationship between branching-diffusion martingales and travelling waves.

2.3 The *Single-Particle Model* and *Many-to-One* theorem

The use of a single-particle model was a key ingredient in the proofs of $\mathcal{L}^p$-convergence of the martingale in Harris and Williams [21], and depended on two important ideas: a many-to-one lemma which reduced expectation calculations to the single-particle model, combined with a change-of-measure technique.

The proof that we give in section 2.5 of the $\mathcal{L}^p$-convergence of the martingale $Z_\lambda(t)$ (for certain p) follow very much in this spirit and therefore we too define a separate single-particle model:

Definition 2.3.1 (The single-particle model) *We suppose that a process* (ξ_t, η_t) *exists on* J *such that under a measure* $\mathbb{P}$ *it behaves stochastically like the branching-diffusion particles of* $\mathbb{X}_t$ *except that no branching occurs. Precisely, this means that* η_t *is an irreducible, time-reversible Markov chain on* I *with Q-matrix* θQ *and invariant measure* $\pi = \{\pi_1, \dots, \pi_n\}$*, whilst* ξ_t *moves as a Brownian motion with zero drift and diffusion coefficient* $a(y) > 0$ *whenever* η_t *is in state* y*:*

$$\mathrm{d}\xi_t = a(\eta_t)^{\frac{1}{2}}\,\mathrm{d}B_t, \tag{2.6}$$

for a $\mathbb{P}$*-Brownian motion* B_t*. We note that the formal generator of this process* (ξ_t, η_t) *is:*

$$\mathcal{H}F(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + \theta\sum_{j\in I} Q(y,j)F(x,j), \qquad (F : J \to \mathbb{R}). \tag{2.7}$$

We suppose that $\mathcal{G}_t$ *is the natural filtration of the single-particle process* (ξ_t, η_t)*, and define* $\mathcal{G}_\infty$ *to be the smallest* σ*-algebra containing each of the* $\mathcal{G}_t$*:*

$$\mathcal{G}_\infty = \sigma\Big(\bigcup_{t\geq 0} \mathcal{G}_t\Big). \tag{2.8}$$

We remark that it is therefore implicit that $\mathbb{P}$ *is a measure defined on* $\mathcal{G}_\infty$*.*

The existence of such a measure $\mathbb{P}$ can be deduced from the standard theorems on the construction of probability measures; in this case we could build $\mathbb{P}$ via the desired finite-dimensional distributions of our single-particle process. For instance, Rogers and Williams [51] contains a full discussion of these matters.

Theorem 2.3.2 (Many-to-One theorem) *For any measurable function* $f : J \to \mathbb{R}$ *we have*

$$P^{x,y}\Big(\sum_{u\in N_t} f(X_u(t), Y_u(t))\Big) = \mathbb{P}^{x,y}\Big(e^{\int_0^t R(\eta_s)\,\mathrm{d}s} f(\xi_t, \eta_t)\Big).$$

For two important reasons we do not prove this Many-to-One theorem immediately, but wait until Section 3.6.3 where we give a spine proof: firstly the spine proof is much easier, and secondly the proof of Harris and Williams (which is also found in Champneys *et al* [4]) is based on resolvents and the Feynman-Kac formula and thus can deal only with functions of the *current* space-type locations of the particles – this is the term $f(X_u(t), Y_u(t))$ in Theorem 2.3.2 – whilst the spine proof can cope the most general functions that may even depend on the complete history of the particles. Our improved version is found at Theorem 3.6.5.

2.3.1 A single-particle martingale and measure change

We now show by direct methods that

$$\zeta_\lambda(t) := v_\lambda(\eta_t) e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}, \tag{2.9}$$

defines a martingale for each $\lambda \in \mathbb{R}$, with respect to the natural filtration $(\mathcal{G}_t)_{t\geq 0}$ for the single-particle model (ξ_t, η_t).

The following proof was actually given by Champneys *et al* [4], and we include it here for completeness with slightly more detail.

Theorem 2.3.3 *For each $\lambda \in \mathbb{R}$, ζ_λ is a martingale.*

Proof: An application of Lemma 2.2.2 with the generator (2.7) shows that ζ_λ is a local martingale, and the aim of the following is to show that $\sup_{s\le t}\zeta_\lambda(s)$ is integrable for each $t>0$, whence it dominates $\zeta_\lambda(s)$ on the time interval $[0,t]$ and the dominated convergence theorem applies to give that ζ_λ is actually a true martingale.

We first note that if $\lambda<0$ then

$$\zeta_{-\lambda}(t) = v_\lambda(\eta_t)e^{\int_0^t R(\eta_s)\,\mathrm{d}s}e^{\lambda(-\xi_t)-E_\lambda t}$$

since $v_{-\lambda}=v_\lambda$ and $E_{-\lambda}=E_\lambda$. Then because $-\xi_t$ has the same law as ξ_t, it is clear that $\zeta_\lambda(t)$ has the same law as $\zeta_{-\lambda}(t)$. Consequently, rather than deal with all $\lambda\in\mathbb{R}$ we need only prove that ζ_λ is a martingale for all $\lambda\le 0$, which we now go on to do.

It is useful to separate ζ_λ as:

$$\zeta_\lambda(t) = W_\lambda(t)\times e^{\lambda\xi_t}, \tag{2.10}$$

where $W_\lambda(t) := v_\lambda(\eta_t)e^{\int_0^t r(\eta_s)\,\mathrm{d}s}e^{-E_\lambda t}$ depends only on the Markov chain. Having done that we note that in the finite-type model, we have an immediate almost sure bound on $W_\lambda(t)$:

$$W_\lambda(t)\le K_1(y,t),\qquad \mathbb{P}^{x,y}\text{-a.s.},\qquad \text{where } K_1(y,t):=\sup_{t\ge 0} v_\lambda(\eta_t)e^{\int_0^t r(\eta_s)\,\mathrm{d}s}e^{-E_\lambda t}. \tag{2.11}$$

Just from this bound and (2.10) it follows that

$$\sup_{s\le t}\zeta_\lambda(s)\le \left\{\begin{array}{cc} K_1(y,t) & \text{if } \lambda=0\\ K_1(y,t)\exp(\lambda\inf_{s\le t}\xi_s) & \text{if } \lambda<0\end{array}\right\}\qquad \mathbb{P}^{x,y}\text{-a.s.} \tag{2.12}$$

The case $\lambda=0$ is therefore done already; we assume from now on that $\lambda<0$.

Because of (2.12), if we can show that $\mathbb{P}^{x,y}\exp(\lambda\inf_{s\le t}\xi_s)$ is finite then it will follow that $\mathbb{P}^{x,y}(\sup_{s\le t}\zeta_\lambda(s))$ is finite. The proof in Champneys *et al* [4] for the two-type model does this by getting an estimate of the left-tail of the distribution function of $\inf_{s\le t}\xi_s$, that clearly shows that $\exp(\lambda\inf_{s\le t}\xi_s)$ is integrable, and we follow this line here.

Doob's supermartingale inequality says that for any $\alpha>0$

$$\mathbb{P}^{x,y}\big(\sup_{s\le t}\zeta_\mu(s)\ge\alpha\big)\le\alpha^{-1}\mathbb{P}^{x,y}\big(\zeta_\mu(0)\big)=\alpha^{-1}v_\mu(y)e^{\mu x}. \tag{2.13}$$

We can use (2.13) to get an estimate on $\mathbb{P}^{x,y}\big(\inf_{s\le t}\zeta_\mu\le -u\big)$, where (for reasons that become clear at (2.18)) we assume that $\mu<\lambda$; simple rearrangements give:

$$\begin{aligned}\mathbb{P}^{x,y}\big(\inf_{s\le t}\xi_s\le -u\big) &= \mathbb{P}^{x,y}\big(e^{\mu\left(\inf_{s\le t}\xi_s\right)}\ge e^{-\mu u}\big)\\ &= \mathbb{P}^{x,y}\big(\exists\tilde{s}\in[0,t]: W_\mu(\tilde{s})e^{\mu\xi_{\tilde{s}}}\ge W_\mu(\tilde{s})e^{-\mu u}\big).\end{aligned} \tag{2.14}$$

It is clear that for any $\tilde{s}\in[0,t]$,

$$W_\mu(\tilde{s})e^{\mu\xi_{\tilde{s}}}\ge W_\mu(\tilde{s})e^{-\mu u}\quad\text{implies}\quad \sup_{s\le t}\zeta_\mu(s)\ge W_\mu(\tilde{s})e^{-\mu u},$$

whence continuing from (2.14) we have

$$\begin{aligned}
\mathbb{P}^{x,y}\big(\inf_{s\le t}\xi_s \le -u\big) &= \mathbb{P}^{x,y}\big(\exists \tilde{s}\in[0,t]: \sup_{s\le t}\zeta_\mu(s) \ge W_\mu(\tilde{s})e^{-\mu u}\big) \\
&\le \mathbb{P}^{x,y}\big(\sup_{s\le t}\zeta_\mu(s) \ge \sup_{s\le t}W_\mu(s)\times e^{-\mu u}\big) \qquad (2.15)\\
&\le \mathbb{P}^{x,y}\big(\sup_{s\le t}\zeta_\mu(s) \ge K_1(y,t)e^{-\mu u}\big) \qquad (2.16)
\end{aligned}$$

We now apply the inequality (2.13) to (2.16) to get

$$\mathbb{P}^{x,y}\big(\inf_{s\le t}\xi_s \le -u\big) \le K_1^{-1}v_\mu(y)e^{\mu x}e^{\mu u} = K_2(x,y,t)e^{\mu u}, \qquad (2.17)$$

where we are defining $K_s(x,y,t) := K_1^{-1}v_\mu(y)e^{\mu x}$. Using this bound and integrating by parts:

$$\begin{aligned}
\mathbb{P}^{x,y}\big(\exp(\lambda\inf_{s\le t}\xi_s)\big) &\le K_3(x,y,t)\int_{-\infty}^{0} e^{\lambda z}\mathbb{P}^{x,y}\big(\inf_{s\le t}\xi_s \in \mathrm{d}z\big) \\
&= K_3(x,y,t)\int_{-\infty}^{0} \lambda e^{\lambda z}\mathbb{P}^{x,y}\big(\inf_{s\le t}\xi_s \le z\big)\mathrm{d}z \\
&\le K_4(x,y,t)\int_{0}^{\infty} e^{(\lambda-\mu)z}\mathrm{d}z \\
&< \infty.
\end{aligned}$$

Returning to (2.12) we see that this shows $\sup_{s\le t}\zeta_\lambda(s) \in \mathcal{L}^1(\mathbb{P}^{x,y})$. We note that from (2.18) it is clear why we chose $\mu < \lambda$. □

Corollary 2.3.4 *$Z_\lambda(t)$ is a true martingale*

Proof: We know that $Z_\lambda(t)$ is a submartingale, and Theorem 2.3.2 gives $\mathbb{P}^{x,y}Z_\lambda(t) = \mathbb{P}^{x,y}\zeta_\lambda(t)$, from which it is immediate that $Z_\lambda(t)$ is in fact a true martingale. □

The idea of changing the measures is absolutely central to the new spine approaches, and the Lyons *et al* papers [40, 34, 41] used a martingale corresponding to Z_λ as the Radon-Nikodym derivative. However, measure-change arguments were also extensively used in the Harris and Williams paper [21]; the difference is that they used the martingale ζ_λ as the derivative, and consequently were dealing with new measures only for the single-particle model. In our new formulation of the spine approach laid out in Chapter 3 we unify these two measure changes. For now, we introduce the new measures only for the single-particle model:

Definition 2.3.5 *For each $\lambda\in\mathbb{R}$ we define a new measure $\mathbb{P}_\lambda$ for the single-particle model,*

$$\frac{\mathrm{d}\mathbb{P}_\lambda}{\mathrm{d}\mathbb{P}} = \frac{\zeta_\lambda(t)}{\zeta_\lambda(0)} \qquad \text{on } \mathcal{G}_t, \qquad (2.18)$$

(the term $\zeta_\lambda(0)$ ensures that $\mathbb{P}_\lambda$ has a mass of 1).

Palmowski and Rolski [48] give an interesting account of such changes of measure – they call them *exponential* changes of measure. Just in case the reader is concerned about questions of the *existence* of such a measure $\mathbb{P}_\lambda$, we have included a brief discussion of the measure-theoretic issues in a paragraph at the end of this section.

We deduce how the process behaves under $\mathbb{P}_\lambda$ by looking at its generator.

Theorem 2.3.6 *The generator of the single-particle model (ξ_t, η_t) under the measure $\mathbb{P}_\lambda$ is*

$$\mathcal{H}_\lambda F(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + a(y)\lambda\frac{\partial F}{\partial x} + \theta\sum_{j\in I} Q_\lambda(y,j)F(x,j), \tag{2.19}$$

where θQ_λ is an honest (stochastic) Q-matrix defined, with $V_\lambda := diag[v_\lambda(1), \ldots, v_\lambda(n)]$, as:

$$\theta Q_\lambda = V_\lambda^{-1}\theta Q V_\lambda + \frac{\lambda^2}{2}A + R - E_\lambda, \tag{2.20}$$

that is,

$$\theta Q_\lambda(i,j) = \begin{cases} \theta Q(i,j)\frac{v_\lambda(j)}{v_\lambda(i)} & \text{if } i \neq j \\ \theta Q(i,i) + \frac{\lambda^2}{2}a(i) - E_\lambda + r(i) & \text{if } i = j. \end{cases}$$

Proof: The proof is based on the standard result for changes of measure given by an h-transform. Details can be found in Rogers and Williams [51]. □

Thus when η_t is in state y, ξ_t is moving like a Brownian motion with diffusion coefficient $a(y)$ and a drift of $\lambda a(\eta_t)$, whilst η_t remains a Markov chain but with the tilted Q-matrix θQ_λ. The fact that θQ is self-adjoint with respect to the inner-product $\langle\cdot,\cdot\rangle_\pi$ means that we can easily identify the invariant measure:

Theorem 2.3.7 *The invariant measure of θQ_λ is given as: $\pi_\lambda(y) = v_\lambda(y)^2\pi(y)$.*

Proof: Because θQ is self-adjoint (see Theorem 2.1.2), we can write $\theta Q = \Pi^{-1}\theta Q^T\Pi$. Now we have, for any vector z

$$\begin{aligned} Q_\lambda z &= V_\lambda^{-1}\left(\frac{\lambda^2}{2}A + \theta Q + R - E_\lambda\right)V_\lambda z \\ &= V_\lambda^{-1}\Pi^{-1}\left(\frac{\lambda^2}{2}A + \theta Q^T + R - E_\lambda\right)\Pi V_\lambda z \\ &= (V_\lambda^2\Pi)^{-1}V\left(\frac{\lambda^2}{2}A + \theta Q^T + R - E_\lambda\right)V_\lambda^{-1}(V_\lambda^2\Pi)z \\ &= (V_\lambda^2\Pi)^{-1}Q_\lambda^T(V_\lambda^2\Pi)z. \end{aligned}$$

We also know $Q_\lambda\mathbf{1} = \mathbf{0}$, and so it must be that $Q_\lambda^T(v_\lambda^2\pi) = \mathbf{0}$, which is to say that $\pi_\lambda \propto v_\lambda^2\pi$. Furthermore, when we chose v_λ, we insisted that $\sum_1^n v_\lambda(i)^2\pi(i) = 1$, and this now forces equality: $\pi_\lambda = v_\lambda^2\pi$. □

Harris and Williams [21] noted that for their branching-diffusion model, changes of measure are particularly useful when combined with the Many-to-One theorem. This is true in general, and we remark that in their specific work they introduced a new *a priori* known martingale for Orstein-Uhlenbeck processes to do this. In fact it can be shown that this apparently different change-of-measure martingale is in fact equal to their single-particle martingale $\zeta_\lambda(t)$, and in general it is the single-particle martingale that should be used as we show below.

The following results are the main reasons for using the change of measure on the single-particle model.

Theorem 2.3.8 *Let f be a measurable function on I and choose a fixed $\lambda \in \mathbb{R}$. Then for all $y \in I$,*

$$P^{0,y}\Big(\sum_{u\in N_t} f(Y_u(t))e^{\lambda X_u(t)-E_\lambda t}\Big) = v_\lambda(y)\,\mathbb{P}_\lambda^{0,y}\Big(\frac{f}{v_\lambda}(\eta_t)\Big). \tag{2.21}$$

Proof: An application of the Many-to-One lemma immediately reduces the expectation to a single-particle calculation,

$$P^{0,y}\Big(\sum_{u\in N_t} f(Y_u(t))e^{\lambda X_u(t)-E_\lambda t}\Big) = \mathbb{P}^{0,y}\Big(f(\eta_t)\,e^{\int_0^t R(\eta_s)\mathrm{d}s}e^{\lambda\xi_t - E_\lambda t}\Big)$$

which can easily be tackled by the change of measure to arrive at

$$\begin{aligned} P^{0,y}\Big(\sum_{u\in N_t} f(Y_u(t))e^{\lambda X_u(t)-E_\lambda t}\Big) &= \mathbb{P}^{0,y}\Big(f(\eta_t)\,e^{\int_0^t R(\eta_s)\mathrm{d}s}e^{\lambda\xi_t - E_\lambda t}\Big) \\ &= v_\lambda(y)\mathbb{P}^{0,y}\Big(f\,\frac{v_\lambda(\eta_0)}{v_\lambda(\eta_t)} \times \frac{\zeta_\lambda(t)}{v_\lambda(y)}\Big) \\ &= v_\lambda(y)\,\mathbb{P}_\lambda^{0,y}\Big(\frac{f}{v_\lambda}(\eta_t)\Big). \end{aligned}$$

□

Corollary 2.3.9 *For each f and each λ there is an upper-bound $K(f,\lambda)$ such that*

$$P^{0,y}\Big(\sum_{u\in N_t} f(Y_u(t))e^{\lambda X_u(t)-E_\lambda t}\Big) < K(f,\lambda)$$

for all $y \in I$ and for all $t \in \mathbb{R}$.

Proof: Since we are working with a finite-state Markov chain, it is trivial that for each $y \in I$,

$$\sup_{t\in\mathbb{R}}\Big[v_\lambda(y)\,\mathbb{P}_\lambda^{0,y}\Big(\frac{f}{v_\lambda}(\eta_t)\Big)\Big] < \infty,$$

and this implies the result. □

A remark on the existence of new measures

To be entirely proper, Definition 2.3.5 should have been phrased as a theorem concerning the existence of a measure $\mathbb{P}_\lambda$ such that equation (2.18) holds. The issue here is that whilst (2.18) can be correctly used to define the measure $\mathbb{P}_\lambda$ on each of the σ-algebras $\mathcal{G}_t$ for each *finite* time

t, it is not possible to therefrom deduce the existence of a measure $\mathbb{P}_\lambda$ on the σ-algebra $\mathcal{G}_\infty$, defined at 2.8 as the smallest σ-algebra containing each of the $\mathcal{G}_t$.

However, Theorem 2.3.6 can be used to resolve the issue in a simple way. At Definition 2.3.1 we supposed the existence of the single-particle measure $\mathbb{P}$ as a probability measure which makes (ξ_t, η_t) a Markov process with generator given at equation (2.7); here we can use exactly the same ideas to directly suppose the existence of the measure $\mathbb{P}_\lambda$ on $\mathcal{G}_\infty$ such that the single-particle process (ξ_t, η_t) has the generator stated at equation (2.19).

2.4 Perron-Frobenius Theory for $\frac{1}{2}\lambda^2 A + \theta Q + R$ and E_λ

In section 2.5 we shall use a generalization of the proof of Champneys *et al* [4] to show that, dependent on λ, the Z_λ martingales are either $\mathcal{L}^p$-bounded or converge to zero. For both the classical and spine approaches the proofs rely importantly on certain properties of the function $\lambda \mapsto E_\lambda$; in the classical proofs the properties of the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$ are also crucial. In this section we consider these features, which follow mostly from a combination of Perron-Frobenius for the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$ and the well-known (Rayleigh-Ritz) representation (2.22) for the largest eigenvalue of a self-adjoint matrix.

In Harris and Williams [21] they were able to explicitly determine the algebraic form of E_λ, from which convexity and other properties followed easily. In general this is difficult to determine, but in our finite-type model we are able to prove convexity in Theorem 2.4.8 and find explicitly the value of the derivative in Theorem 2.4.9. Then, in Theorem 2.4.11, we prove the existence of a global minimum for the associated function c_λ, which determines the minimum possible speed of travelling waves in later chapters.

Some of the analysis and matrix results of Champneys *et al* were generalized to deal with the n-type case in a paper by Crooks [11] which also relies on Perron-Frobenius theory and presents alternative proofs to our Theorem 2.4.8 and the differentiability property in Theorem 2.4.9.

The main reference is Seneta [53], where we are interested in his Perron-Frobenius theory for the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$. We are assuming the type Markov chain to be irreducible, which according to Norris [45] (page 111, Theorem 3.2.1) can be taken to be defined as

Definition 2.4.1 *A (continuous-time) Markov chain with Q-matrix $Q = [q_{ij}]$ is said to be* irreducible *if and only if for each pair i, j of states,*

$$q_{i_0 i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} > 0$$

for some states $i_0, i_1, \ldots, i_n$ with $i_0 = i$ and $i_n = j$.

On the other hand, Seneta gives an alternative definition of irreducibility, first only for non-negative matrices:

Definition 2.4.2 *An $n \times n$ non-negative matrix $T = [t_{ij}]$ is* irreducible *if for every pair i, j of its indices there exists a positive integer $m = m(i, j)$ such that $t_{ij}^{(m)} > 0$, where $[t_{ij}^{(m)}] := T^m$.*

To give a definition of irreducibility to matrices like $\frac{1}{2}\lambda^2 A + \theta Q + R$, which are non-negative off the diagonal but possibly contain negative values on the diagonal (he calls these 'ML matrices'), Seneta observes that for any given ML matrix B, we can always find a $\mu \in \mathbb{R}$ such that

$$T := \mu I + B$$

is a non-negative matrix. This allows him (on page 46) to define irreducibility for an ML matrix:

Definition 2.4.3 *An ML matrix B is called* irreducible *if the matrix T is irreducible, for some μ large enough to make T non-negative.*

As a matter of fact, it is not difficult to show that these alternative definitions 2.4.1, 2.4.2 and 2.4.3 are equivalent (this is Theorem 1.2.1 in Norris [45]), and we can therefore apply Seneta's Perron-Frobenius theorem for irreducible ML matrices (Theorem 2.6, page 46) to $\frac{1}{2}\lambda^2 A + \theta Q + R$. We mention only the results that interest us and refer the reader to Seneta [53] for the details and proofs:

Theorem 2.4.4 *The matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$ has a real eigenvalue E_λ (which we will refer to as* the Perron-Frobenius eigenvalue*) such that $E_\lambda > Re\, \tau$ for any other eigenvalue τ. With E_λ is associated a* strictly positive *right-eigenvector v_λ unique up to constant multiples, and furthermore E_λ is a simple root of the characteristic equation.*

We always suppose that v_λ is normalized with $\langle v_\lambda, v_\lambda \rangle_\pi = 1$.

Many positivity properties of a non-negative, irreducible matrix relate closely to the size of its Perron-Frobenius eigenvalue: in Section 2.5.2 we will need to show that a certain matrix of the form $I - T$ (where T is non-negative irreducible) has a strictly non-negative inverse, and a result from Perron-Frobenius theory can deal with this. In Seneta this is Theorem 2.1 page 30 and Corollaries 1 and 2, which here we present as a single theorem:

Theorem 2.4.5 *If T is a non-negative, irreducible matrix and $\Lambda_{PF}(T) < 1$, where $\Lambda_{PF}(T)$ is the Perron-Frobenius eigenvalue of T, then the inverse $(I - T)^{-1}$ exists and is strictly positive – ie. all its elements are strictly positive.*

When we use this theorem we shall also need the following monotonicity property that holds for non-negative irreducible matrices – part (e) of Seneta's Theorem 1.5:

Theorem 2.4.6 *If T is a non-negative, irreducible matrix and $0 \leq B \lneqq T$, then each eigenvalue β of B satisfies $\beta < \Lambda_{PF}(T)$.*

Remark: We are using an ordering notation for matrices found in Seneta: for a matrix $T = [t_{ij}]$ he writes $T \geq 0$ if and only if $t_{ij} \geq 0$ for all i, j; if it is further true that $t_{ij} > 0$ for *all* i, j then he writes $T > 0$. Above we use an extra notation $T \gneqq 0$ to mean $t_{ij} \geq 0$ for all i, j *but with $t_{ij} > 0$ for at least one choice of i and j* – guaranteeing that $T \neq 0$. This notation extends in an obvious way to expressions such as $B \leq T$, for example:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} < \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \leq \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \lneqq \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}.$$

The eigenvalue E_λ

Turning our attention to E_λ, the self-adjoint property of $\frac{1}{2}\lambda^2 A + \theta Q + R$ gives an alternative representation from which important properties follow.

Theorem 2.4.7

$$E_\lambda = \sup_{\|v\|_\pi = 1} \left\langle \left((\lambda^2/2)A + \theta Q + R\right) v, v \right\rangle_\pi, \tag{2.22}$$

since it is the rightmost eigenvalue (see Kreyzig [33]).

This supremum is attained at the eigenvector v_λ:

$$E_\lambda = \left\langle \left((\lambda^2/2)A + \theta Q + R\right) v_\lambda, v_\lambda \right\rangle_\pi, \tag{2.23}$$

and we can now show that E_λ is a strictly convex function of λ throughout $\mathbb{R}$:

Theorem 2.4.8 *$\lambda \mapsto E_\lambda$ is strictly convex on $\mathbb{R}$.*

Proof: Suppose that $\lambda_1, \lambda_2 \in \mathbb{R}$ are distinct, and $0 < p, q < 1$ such that $p + q = 1$. We wish to show that $E(p\lambda_1 + q\lambda_2) < pE(\lambda_1) + qE(\lambda_2)$.

The following inequalities hold:

$$\begin{aligned}
E(p\lambda_1 + q\lambda_2) &= \left\langle \left(\frac{(p\lambda_1 + q\lambda_2)^2}{2} A + \theta Q + R \right) v_{p\lambda_1 + q\lambda_2}, v_{p\lambda_1 + q\lambda_2} \right\rangle_\pi \\
&< p\left\langle \left(\frac{\lambda_1^2}{2} A + \theta Q + R \right) v_{p\lambda_1 + q\lambda_2}, v_{p\lambda_1 + q\lambda_2} \right\rangle_\pi + \\
&\quad q\left\langle \left(\frac{\lambda_2^2}{2} A + \theta Q + R \right) v_{p\lambda_1 + q\lambda_2}, v_{p\lambda_1 + q\lambda_2} \right\rangle_\pi \\
&\leq p\left\langle \left(\frac{\lambda_1^2}{2} A + \theta Q + R \right) v_{\lambda_1}, v_{\lambda_1} \right\rangle_\pi + \\
&\quad q\left\langle \left(\frac{\lambda_2^2}{2} A + \theta Q + R \right) v_{\lambda_2}, v_{\lambda_2} \right\rangle_\pi \\
&= pE(\lambda_1) + qE(\lambda_2).
\end{aligned}$$

firstly because $\lambda \mapsto \lambda^2/2$ is strictly convex and A is positive definite, and also because the two inner products attain their maximum at v_{λ_1} and v_{λ_2} respectively. Note that the first inequality would not be strict if it happened that $v_{p\lambda_1+q\lambda_2} = 0$, but this is excluded by the Perron-Frobenius theorem which states that the eigenvector is strictly positive. $\square$

Remark: The above proof depends on the representation of E_λ given by (2.22), but we note that a paper by Cohen [10] shows that the convexity of E_λ would hold even if the matrix was not self-adjoint – Cohen shows that the lead eigenvalue of all so-called *essentially non-negative* matrices (which Senata calls ML matrices) is always a convex function of the diagonal elements. This is also the approach found in Crooks [11], Lemma 3.7 page 23, which in fact uses this result by Cohen.

Another simple use of the representation (2.22) is to see that $E(0) > 0$: (below $\mathbf{1}$ is the vector with all entries 1)

$$\begin{aligned} E(0) &= \sup_{\|v\|_\pi = 1} \left\langle (\theta Q + R)\, v, v \right\rangle_\pi \\ &\geq \left\langle (\theta Q + R)\, \mathbf{1}, \mathbf{1} \right\rangle_\pi \\ &= \sum_{k=1}^{n} r_k \pi_k > 0. \end{aligned} \tag{2.24}$$

Being a convex function, we immediately know that the left- and right-derivatives $E'_-(\lambda)$ and $E'_+(\lambda)$ exist everywhere, and that in fact they must be equal *almost everywhere* with respect to Lebesgue measure on $\mathbb{R}$ (see Tiel [54], Theorem 1.6 page 4). Actually, since we are dealing with a finite-dimensional vector space, the Perron-Frobenius theorem 2.4.4 can be used which states that $E(\lambda)$ is a simple root of the characteristic equation, and it therefore follows from the Implicit Function Theorem that it is actually a C^∞ function of λ – i.e. the left- and right-derivatives are equal *everywhere.* Because of the representation (2.22), calculating the derivative is also quite straightforward.

Theorem 2.4.9 *The eigenvalue $E(\lambda)$ is infinitely differentiable on $\mathbb{R}$, and its derivative is given by:*

$$E'(\lambda) = \lambda \left\langle A v_\lambda, v_\lambda \right\rangle_\pi. \tag{2.25}$$

Proof: Because the eigenvalues of $(\lambda^2/2)A + \theta Q + R$ are just the zeroes of a polynomial (the characteristic equation) whose coefficients depend continuously on λ, it follows that the eigenvalues are (branches of) analytic functions possibly having analytic singularities (see pages 63-64 of Kato [27]). However Theorem 2.4.4 states that for any real λ the rightmost eigenvalue $E(\lambda)$ is always finite, and is a simple root of the characteristic equation. Thus $E(\lambda)$ must actually be analytic, and therefore is an infinitely differentiable function of $\lambda \in \mathbb{R}$.

If we use the shorthand $f_\lambda(\cdot) = \left\langle ((\lambda^2/2)A + \theta Q + R)\cdot, \cdot \right\rangle_\pi$ then

$$\begin{aligned} E(\lambda + \varepsilon) - E(\lambda) &= f_{\lambda+\varepsilon}(v_{\lambda+\varepsilon}) - f_\lambda(v_\lambda) \\ &\geq f_{\lambda+\varepsilon}(v_\lambda) - f_\lambda(v_\lambda) \\ &= \lambda\varepsilon \left\langle A v_\lambda, v_\lambda \right\rangle_\pi + \frac{\varepsilon^2}{2} \left\langle A v_\lambda, v_\lambda \right\rangle_\pi. \end{aligned}$$

It therefore follows that

$$E'_+(\lambda) \geq \lambda \left\langle A v_\lambda, v_\lambda \right\rangle_\pi.$$

Similarly,

$$\begin{aligned} E(\lambda) - E(\lambda - \varepsilon) &= f_\lambda(v_\lambda) - f_{\lambda-\varepsilon}(v_{\lambda-\varepsilon}) \\ &\leq f_\lambda(v_\lambda) - f_{\lambda-\varepsilon}(v_\lambda) \\ &= \lambda\varepsilon \left\langle A v_\lambda, v_\lambda \right\rangle_\pi + \frac{\varepsilon^2}{2} \left\langle A v_\lambda, v_\lambda \right\rangle_\pi. \end{aligned}$$

which gives

$$E'_-(\lambda) \leq \lambda \left\langle A v_\lambda, v_\lambda \right\rangle_\pi.$$

Since we have already shown that E is differentiable, it follows that $E'_+(\lambda) = E'_-(\lambda)$ and so $E'(\lambda) = \lambda\langle Av_\lambda, v_\lambda\rangle_\pi$.

□

Definition 2.4.10 *We define the speed function* $c : (-\infty, 0) \to \mathbb{R}$ *as:*

$$c_\lambda := \frac{E_\lambda}{-\lambda}. \tag{2.26}$$

We refer to the function c_λ as the speed function since it relates to the asymptotic speed of the travelling waves associated with the martingale $Z_\lambda(t)$; see section 2.6 for details of the relationship between branching-diffusion martingales and travelling waves.

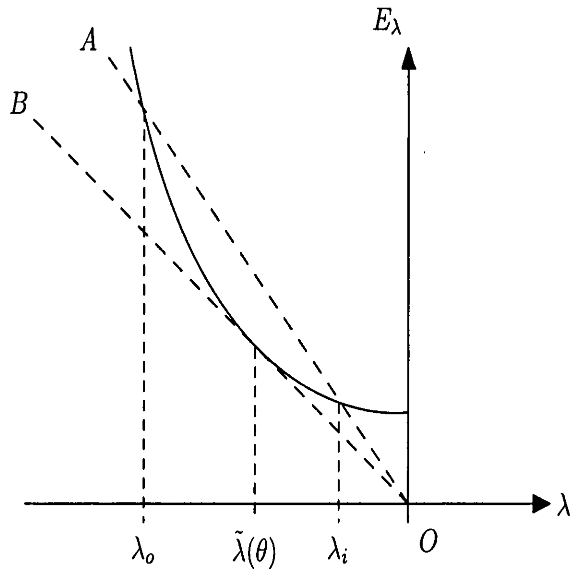
Throughout the thesis we only allow the parameter λ to vary over $(-\infty, 0]$, but by simple arguments of symmetry all results can be generalized to cover the whole of $\mathbb{R}$ – it is only a question of convenience that we do not consider positive λ. In particular, if we write

$$\ldots \textit{if and only if } \lambda \in (\tilde{\lambda}(\theta), 0],$$

it is to be understood as meaning

$$\ldots \textit{if and only if } \lambda \in \big(-|\tilde{\lambda}(\theta)|, +|\tilde{\lambda}(\theta)|\big).$$

The figure below is a schematic picture of the graph of E_λ, where we can see c_λ geometrically as (-1 times) the gradient of the ray connecting the origin O to the point (λ, E_λ) on the curve. We have drawn in two such rays OA and OB.



Because of the geometric interpretation, it is immediate that $c_{\lambda_o} = c_{\lambda_i}$, since they both share the same ray OA, and that on the negative half-line c_λ reaches a minimum at the point $\tilde{\lambda}(\theta)$ where the ray OB becomes a tangent to E_λ. This is the content of the following theorem:

Theorem 2.4.11 *On $(-\infty, 0)$ the function c_λ is differentiable, has just one minimum at a single point $\tilde{\lambda}(\theta)$, and is strictly increasing to $+\infty$ as $\lambda \downarrow -\infty$ or as $\lambda \uparrow 0$.*

Proof: As $\lambda \uparrow 0$ it is clear that $c_\lambda \to \infty$, because $E(0) > 0$ as shown at (2.24). On the other hand, since $E(\lambda)$ can be written as a supremum, it follows that

$$E(\lambda) \geq \left\langle \left((\lambda^2/2)A + \theta Q + R \right) \mathbf{1}, \mathbf{1} \right\rangle_\pi = \sum_{k=1}^{n} \left(\frac{\lambda^2}{2} a_k + r_k \right) \pi_k,$$

where $\mathbf{1}$ is the vector with all entries 1 (trivially $\|\mathbf{1}\|_\pi = 1$). Consequently,

$$c_\lambda \geq \sum_{k=1}^{n} \left(\frac{-\lambda}{2} a_k + \frac{r_k}{\lambda} \right) \pi_k,$$

showing that $c_\lambda \to \infty$ as $\lambda \downarrow -\infty$. Therefore c_λ has at least one minimum on $(-\infty, 0)$.

Since $E(\lambda)$ is differentiable it follows that c_λ is too, and

$$c'_\lambda = \frac{\lambda E'(\lambda) - E(\lambda)}{\lambda^2} =: \frac{f(\lambda)}{\lambda^2}.$$

Then $f'(\lambda) = \lambda E''(\lambda) < 0$ on $(-\infty, 0)$, since $E(\lambda)$ is strictly convex on $(-\infty, 0)$. Consequently, f is strictly decreasing on $(-\infty, 0)$ and since $f(0) = -E(0) < 0$ we conclude that $f(\lambda)$, and therefore c'_λ, has exactly one zero on $(-\infty, 0)$.

Thus c_λ has exactly one minimum on $(-\infty, 0)$, when $c_{\tilde{\lambda}(\theta)} = -E'(\tilde{\lambda}(\theta))$.

□

Definition 2.4.12 *We give the name $\tilde{c}(\theta)$ to the value of the minimum of c_λ on the negative half-line:*

$$\tilde{c}(\theta) := \inf_{\lambda < 0} c_\lambda = c_{\tilde{\lambda}(\theta)}.$$

For later proofs, the following corollary importantly lays out the behaviour of the eigenvalue E_λ:

Corollary 2.4.13 *For each $\lambda \in (\tilde{\lambda}(\theta), 0]$ there is some $p > 1$ such that $pE_\lambda - E_{p\lambda} > 0$; on the other hand, if $\lambda < \tilde{\lambda}(\theta)$ there is no such $p > 1$.*

Proof: If $\lambda = 0$, then $pE_0 - E_0 > 0$ for any $p \in (1, 2]$, since $E_0 \neq 0$. Otherwise, the fact that

$$pE_\lambda - E_{p\lambda} = p\lambda(c_{p\lambda} - c_\lambda)$$

together with the properties of c_λ given in Theorem 2.4.11 imply the result. □

2.5 A first proof of $\mathcal{L}^p$ convergence

Having covered the essentials of Perron-Frobenius theory and the single-particle model, we can prove the following theorem which explicitly shows how the further convergence properties of the branching-diffusion martingales $Z_\lambda(t)$ are dependent on the behaviour of the lead eigenvalue E_λ. We shall later give spine proofs of this same theorem that are significantly better, but include the following 'classical' proofs from our earlier work for comparison and completeness.

We remark that these following proofs of both parts are closely based on the inequalities and ideas used by Champneys *et al* [4] (and also are similar to the proofs of Harris and Williams [21]), but where the Champneys *et al* proofs rely on specific algebraic and matrix results for their explicit functions and matrices, here we have had to formulate proofs using the more general Perron-Frobenius theory laid out in section 2.4.

Theorem 2.5.1 *We remember that for each $\lambda \in \mathbb{R}$, $Z_\lambda(t)$ defines a strictly positive martingale which under P converges a.s. to a finite limit $Z_\lambda(\infty)$.*

1. *If $\lambda \in (\tilde{\lambda}(\theta), 0]$, then there is some $p \in (1,2]$ such that $pE_\lambda - E_{p\lambda} > 0$, and for this same p it follows that $Z_\lambda(t)$ is convergent in $\mathcal{L}^p$.*

2. *Conversely, if $\lambda < \tilde{\lambda}(\theta)$ then $Z_\lambda(t) \to 0$ a.s.*

Note that throughout this thesis we do not deal with the critical case $\lambda = \tilde{\lambda}(\theta)$ since this requires different techniques using 'derivative' martingales – see Kyprianou [35] or Harris [23] for examples in the case of branching Brownian motion.

2.5.1 Part 1: $\lambda \in (\tilde{\lambda}(\theta), 0]$

By Doob's theorem it is only necessary to show that Z_λ is *bounded* in $\mathcal{L}^p$ on order to deduce that it is convergent in $\mathcal{L}^p$.

The conditional form of Jensen's inequality implies that Z_λ^p is a submartingale if $p \in (1,2]$, and since $t \mapsto P^{x,y}(Z_\lambda(t)^p)$ is therefore non-decreasing and always finite, the only way to have *unboundedness* in $\mathcal{L}^p$ is to have

$$\lim_{t \to \infty} P^{x,y}(Z_\lambda(t)^p) = \infty. \tag{2.27}$$

Consequently, we will show that the limit (2.27) cannot be infinite. The following lemma used by Champneys *et al* [4] was originally stated by Neveu [44] and is the important starting point for this:

Lemma 2.5.2 *Let $p \in (1,2]$. For any finite sequence $W_1, \ldots, W_n$ of non-negative independent variables in $\mathcal{L}^p$ and any sequence $l_1, \ldots, l_n$ of non-negative real numbers, we have*

$$\psi\left(\sum_{k=1}^{n} l_k W_k\right) \leq \sum_{k=1}^{n} l_k^p \psi(W_k),$$

where $\psi(W) := P(W^p) - P(W)^p$ for $W \in \mathcal{L}^p$.

We make the usual branching decomposition:

$$Z_\lambda(s+t) = \sum_{u\in N_s} \exp\{\lambda X_u(s) - E_\lambda s\} W^{0,y_u}(t)$$

where $W^{0,y_u}(t)$ behaves like the branching process with one particle started at $(0, y_u)$ where $y_u = Y_u(s)$. Applying the above lemma conditionally on $\mathcal{F}_s$ gives:

$$\begin{aligned}
P^{x,y}\big(Z_\lambda(s+t)^p|\mathcal{F}_s\big) - Z_\lambda(s)^p &\le \sum_{u\in N_s} e^{\lambda p X_u(s) - pE_\lambda s}\big\{P\big(W^{0,y_u}(t)^p|\mathcal{F}_s\big) - P\big(W^{0,y_u}(t)|\mathcal{F}_s\big)^p\big\} \\
&= \sum_{u\in N_s} e^{\lambda p X_u(s) - pE_\lambda s}\big\{P^{0,y_u}(Z_\lambda(t)^p) - P^{0,Y_u(s)}(Z_\lambda(t))^p\big\} \\
&\le \sum_{u\in N_s} e^{\lambda p X_u(s) - pE_\lambda s}\, P^{0,y_u}(Z_\lambda(t)^p). \\
&= \Big(\sum_{u\in N_s} f_t\big(Y_u(s)\big) e^{\lambda p X_u(s) - E_{p\lambda} s}\Big) e^{-(pE_\lambda - E_{p\lambda})s} \qquad (2.28)
\end{aligned}$$

where $f_t(y) = P^{0,y}(Z_\lambda(t)^p)$.

In section 2.3.1 we considered expressions like the bracketed term above, and from Corollary 2.3.9 we can deduce that for each fixed t there is some $K \in \mathbb{R}$ such that

$$P^{x,y}\Big(\sum_{u\in N_s} f_t\big(Y_u(s)\big) e^{\lambda p X_u(s) - E_{p\lambda} s}\Big) < K, \quad \text{for all } s \in [0,\infty).$$

Thus if we fix $t > 0$ and take expectations of (2.28),

$$\begin{aligned}
P^{x,y}(Z_\lambda(s+t)^p) - P^{x,y}(Z_\lambda(s)^p) &\le P^{x,y}\Big(\sum_{u\in N_s} f_t\big(Y_u(s)\big) e^{\lambda p X_u(s) - E_{p\lambda} s}\Big) e^{-(pE_\lambda - E_{p\lambda})s} \\
&\le K e^{-(pE_\lambda - E_{p\lambda})s}.
\end{aligned}$$

At this point it is clear that the behaviour of $pE_\lambda - E_{p\lambda}$ will play a significant part in the rest of the proof. As explained in Corollary 2.4.13:

- If $\lambda = 0$, then

  $$pE_0 - E_0 > 0$$

  for any $p \in (1, 2]$, since $E_0 \ne 0$.

- If $\lambda \in (\tilde\lambda(\theta), 0)$, then for p sufficiently close to 1

  $$pE_\lambda - E_{p\lambda} = -p\lambda(c_\lambda - c_{p\lambda}) > 0,$$

  since $\mu \mapsto c_\mu$ is strictly increasing on $(\tilde\lambda(\theta), 0)$ as we proved at Theorem 2.4.11.

In either case, for any $\lambda \in (\tilde\lambda(\theta), 0]$ we can find a $p > 1$ such that for any $t > 0$ and $s > 0$ and some number $\kappa > 0$:

$$P^{x,y}(Z_\lambda(s+t)^p) - P^{x,y}(Z_\lambda(s)^p) \le K \exp(-\kappa s). \qquad (2.29)$$

This result is telling us that as $s \to \infty$, over a time interval $[s, s+t]$ our Z_λ^p changes by ever decreasing amounts that decay at some exponential rate. The idea of Champneys *et al* is to now string these intervals together:

$$P^{x,y}\big(Z_\lambda((n+1)s+t)^p - Z_\lambda(s+t)^p\big) = \sum_{m\le n} P^{x,y}\big(Z_\lambda((m+1)s+t)^p - Z_\lambda(ms+t)^p\big) \quad (2.30)$$

and each of the terms in this sum can be bounded using (2.29) to give

$$P^{x,y}\big(Z_\lambda((m+1)s+t)^p - Z_\lambda(ms+t)^p\big) \le K' e^{-\kappa t} \exp(-\kappa ms).$$

Thus

$$\sum_{m\le n} P^{x,y}\big(Z_\lambda((m+1)s+t)^p - Z_\lambda(ms+t)^p\big) \le K' e^{-\kappa t} \sum_{m\le n} \exp(-\kappa ms)$$

The sum on the RHS converges as $n \to \infty$, and so choosing $t = s = 1$ in (2.30) we arrive at:

$$P^{x,y}(Z_\lambda(n+2)^p) \le \hat{K}, \qquad \text{for all large } n,$$

hence (2.27) is impossible, and so Z_λ is bounded in $\mathcal{L}^p$.

□

Corollary 2.5.3 *The martingale Z_λ is uniformly integrable whenever $\lambda \in (\tilde{\lambda}(\theta), 0]$.*

Proof: Since the condition $pE_\lambda - E_{p\lambda}$ holds for some $p > 1$ whenever $\lambda \in (\tilde{\lambda}(\theta), 0]$, it follows that the martingale Z_λ is $\mathcal{L}^1$ convergent whenever $\lambda \in (\tilde{\lambda}(\theta), 0]$. This implies the equality:

$$P^{x,y} Z_\lambda(\infty) = P^{x,y} Z_\lambda(0) = e^{\lambda x} v_\lambda(y). \quad (2.31)$$

At the same time, the branching decomposition says that

$$P\big(Z_\lambda(\infty)|\mathcal{F}_s\big) = \sum_{u\in N_s} e^{\lambda X_u(s) - E_\lambda s}\ P^{0,Y_u(s)} Z_\lambda^{s,u}(\infty),$$

where, conditional on $\mathcal{F}_s$, $Z_\lambda^{s,u}(\infty)$ are IID copies of the original $Z_\lambda(\infty)$. Consequently, from the equality (2.31) we have $P^{0,Y_u(s)} Z_\lambda^{s,u}(\infty) = v_\lambda\big(Y_u(s)\big)$ and so

$$P\big(Z_\lambda(\infty)|\mathcal{F}_s\big) = \sum_{u\in N_s} e^{\lambda X_u(s) - E_\lambda s} v_\lambda(Y_u(s)) = Z_\lambda(s),$$

which proves the UI property.

□

2.5.2 Part 2: $\lambda < \tilde{\lambda}(\theta)$

We aim to show that

if $\lambda < \tilde{\lambda}(\theta)$ then $Z_\lambda(t) \to 0$ a.s.

Again our proof is very closely based on that of Champneys *et al* [4], but here we need to make more use of Perron-Frobenius theory of the matrix $\frac{1}{2}\lambda^2 A + \theta Q + R$ stated in section 2.4 to achieve the result.

The following useful inequality was stated in Neveu [44]:

Proposition 2.5.4 *If* $0 < p < 1$ *and* $u, v > 0$ *then* $(u+v)^p \leq u^p + v^p$.

We choose and fix $p \in (0,1)$ and let J be the first jump time of Y_1, and let T be the first branch time of the branching diffusion. Just as in the Champneys *et al* proof, the decomposition:

$$Z_\lambda(\infty) = \begin{cases} \exp\{\lambda[X_1(J) + c_\lambda J]\} Z_\lambda^{(1)}(\infty) & \text{if } J < T, \\ \exp\{\lambda[X_1(T) + c_\lambda T]\}[Z_\lambda^{(2)}(\infty) + Z_\lambda^{(3)}(\infty)] & \text{if } J > T, \end{cases}$$

together with the inequality from the above proposition leads to

$$\begin{aligned} g(y) &:= P^{0,y}(Z_\lambda(\infty)^p) \\ &\leq P^{0,y}\big(e^{\alpha[X_1(J)+c_\lambda J]}\mathbf{1}_{(J<T)}\, g(Y_1(J))\big) + 2 \times \mathbb{P}^{0,y}\big(e^{\alpha[X_1(T)+c_\lambda T]}\mathbf{1}_{(T<J)}\, g(Y_1(T))\big), \end{aligned}$$

where $\alpha := \lambda p$. As Champneys *et al* state, these expectations can be evaluated to give

$$g(y) \leq \frac{\theta \sum_{z \neq y} Q(y,z) g(z) + 2r(y)g(y)}{-\frac{1}{2}\alpha^2 a(y) - \alpha c_\lambda + r(y) + \theta q(y)}, \tag{2.32}$$

the denominator being positive for p near 1; furthermore, since Z_λ itself is positive we immediately deduce that $g \geq 0$ on I. (2.32) can be arranged to give

$$0 \leq \left(\frac{\alpha^2}{2}A + \theta Q + R + \alpha c_\lambda I\right) g = (M_\alpha + \alpha c_\lambda) g, \tag{2.33}$$

where we define

$$M_\alpha := \frac{\alpha^2}{2}A + \theta Q + R,$$

and we now use Perron-Frobenius ideas to show that in fact $g = 0$, and at this point our proof departs from that given by Champneys *et al* since they were able to rely on specific algebraic relationships to obtain this.

First of all we decompose $M_\alpha + \alpha c_\lambda$ as follows:

$$M_\alpha + \alpha c_\lambda = -D_{\alpha,\lambda} + E + F$$

where $D_{\alpha,\lambda}$ is a diagonal matrix,

$$D_{\alpha,\lambda} := \operatorname{diag}\left[-(\frac{\alpha^2}{2}a_i - \theta q_i + r_i + \alpha c_\lambda)\right]_{i=1}^n$$

and E and F are respectively upper and lower diagonal matrices with zeroes on the diagonal (note that this means that $E + F$ is just θQ with zeroes instead of its diagonal elements).

Since v_α is the eigenvector, we know that $(M_\alpha + \alpha c_\alpha)v_\alpha = 0$, and since v_α is a strictly positive vector, it follows from this that

$$\frac{\alpha^2}{2}a_i - \theta q_i + r_i + \alpha c_\alpha < 0, \qquad i = 1, \ldots, n.$$

Therefore, by choosing α close enough to λ (ie. by choosing p close to 1) we can guarantee that the diagonal elements of the matrix $D_{\alpha,\lambda}$ are all strictly positive, whence $D_{\alpha,\lambda}^{-1}$ exists and we can write

$$M_\alpha + \alpha c_\lambda = -D_{\alpha,\lambda}(I - D_{\alpha,\lambda}^{-1}(E+F)).$$

We now aim to show that the matrix $(I - D_{\alpha,\lambda}^{-1}(E+F))^{-1}$ is strictly positive, so that $M_\alpha + \alpha c_\lambda$ must be *strictly negative* – which together with (2.33) will force $g = 0$.

Since $E+F$ is just θQ with zeroes along the diagonal, it inherits its irreducibility from θQ – see definitions 2.4.1 or 2.4.3. Therefore $D_{\alpha,\lambda}^{-1}(E+F)$ is non-negative and irreducible, and from Theorem 2.4.5, it follows that the matrix $(I - D_{\alpha,\lambda}^{-1}(E+F))^{-1}$ exists and is strictly positive if $\Lambda_{PF}(D_{\alpha,\lambda}^{-1}(E+F)) < 1$. The fact that

$$\Lambda_{PF}(D_{\alpha,\alpha}^{-1}(E+F)) = 1$$

is an algebraic consequence of the known result $\Lambda_{PF}(M_\alpha + \alpha c_\alpha) = 0$. Also, because $c_\lambda > c_\alpha$ it follows that

$$D_{\alpha,\lambda}^{-1}(E+F) \lneqq D_{\alpha,\alpha}^{-1}(E+F).$$

Consequently, because of the monotonicity property stated in Theorem 2.4.6 we deduce that

$$\Lambda_{PF}(D_{\alpha,\lambda}^{-1}(E+F)) < 1$$

and hence

$$(I - D_{\alpha,\lambda}^{-1}(E+F))^{-1} > 0.$$

This completes the proof that $g = 0$, whence $P^{0,y}(Z_\lambda(\infty)^p) = 0$. With the 0-1 law stated in Theorem 2.2.3 this implies that $Z_\lambda(\infty) = 0$ almost-surely. □

Remark: Crooks [11] gives an alternative proof that $g = 0$ based on a theorem by Krasnosel'skii [32].

2.6 Travelling waves and FKPP problems

We remark that the work in this section is offered as an adjunct to the main themes of the thesis, to give context and show some new results that were obtained as part of earlier work on the finite-type model. The notation is minimally different from that used elsewhere in this thesis, but is entirely consistent with that used by Champneys *et al* [4] and Harris [23], on which this body of work is based.

2.6.1 The FKPP equation

The Fisher-KPP (FKPP) equation

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial t^2} + f(u), \tag{2.34}$$

has been extensively studied by both analytic and probabilistic techniques.

The 1936 paper by Kolmogorov, Petrovski and Piscounov [31] gave some important results on the FKPP equation, among which:

1. with simple conditions on f, they showed that from arbitrary bounded initial conditions $0 \leq u(0,x) \leq 1$, the equation (2.34) will always develop a solution $u(t,x)$ that remains bounded for all time: $0 \leq u(t,x) \leq 1$ for all $t > 0$.

2. In the long-term, solutions with Heaviside initial conditions will progress along $\mathbb{R}$ at a fixed speed c and will take on a limit shape: $u(t, x + ct) \to v(x)$, where v satisfies

$$c\frac{\mathrm{d}v}{\mathrm{d}x} = k\frac{\mathrm{d}^2 v}{\mathrm{d}x^2} + f(v),$$

and vanishes for $x = -\infty$ and approaches 1 for $x = +\infty$ – in other words, these solutions converge to the *travelling waves solution which has speed c.*

Remark: In the sequel, whenever we refer to travelling waves, we will implicitly suppose that the wave vanishes at $x = -\infty$ and approaches 1 for $x = +\infty$, though we use the term *monotone travelling waves* to emphasize this asymptotic behaviour when it needs to be made clear. This is not to say that other types of non-monotonic travelling waves do not exist or are not interesting, but just that they do not form the object of this work.

2.6.2 Relation to BBM and the McKean Representation

One particular example of the FKPP equation with an f that satisfies the conditions (1) imposed by KPP is the following:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + u^2 - u, \tag{2.35}$$

and its connection with Branching Brownian Motion (BBM) was first laid out in the 1976 paper by McKean [42]. In much the same way that solutions to the heat equation can be expressed as an expectation of a single Brownian motion, solutions $u(t,x)$ of (2.35) can be represented by expectations taken over all the particles in a BBM: if we suppose that there are $N(t)$ particles alive in the BBM at time t, and that there spatial positions are $X_1(t), \ldots, X_{N(t)}(t)$, then the *McKean representation* is just an expectation of a product over the particles –

$$\begin{aligned} u(t,x) &= \mathbb{E}^x\{g(X_1(t)) \cdots g(X_{N(t)}(t))\} \\ &= \mathbb{E}^0\{g(X_1(t) + x) \cdots g(X_{N(t)}(t) + x)\}, \end{aligned}$$

where $u(0,x) = g(x)$ are the initial conditions, and under the measure $\mathbb{P}^x$ (with expectation operator $\mathbb{E}^x$) the BBM starts with a single particle located at x.

With the McKean representation established we can immediately see the probabilistic significance of the Heaviside initial conditions: if

$$g(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$$

then

$$\begin{aligned} u(t,x) &= \mathbb{P}^0(X_1(t) > -x, \ldots, X_{N(t)}(t) > -x) \\ &= \mathbb{P}^0(L_t > -x) \end{aligned}$$

where L_t is defined to be the position of the *leftmost* particle in the BBM at time t. Consequently, when combined with the second KPP result above, the McKean representation proves:

Theorem 2.6.1 *If L_t is defined as the position of the* leftmost *particle in the BBM at time t, then*

$$u(t,x) := \mathbb{P}^0(L_t > -x)$$

is a solution to the FKPP equation (2.35) with Heaviside initial conditions, and

$$\lim_{t\to\infty} \mathbb{P}^0(L_t > -x - \sqrt{2}t) = w_{\sqrt{2}}(x)$$

where $w_{\sqrt{2}}$ is a (monotone) travelling wave solution of (2.35) of speed $\sqrt{2}$ – ie. $w_{\sqrt{2}}$ is a solution of the travelling wave equation

$$0 = \frac{1}{2}w_c'' + cw_c' + w_c^2 - w_c, \tag{2.36}$$

with speed $c = \sqrt{2}$.

The BBM approach not only allowed McKean to offer a simplified proof of the KPP results, it also gave improvements in various directions. For the purposes of this thesis, the following is particularly relevant:

Theorem 2.6.2 *If, for fixed $0 < b < \sqrt{2}$, the initial conditions $g(x)$ satisfy the asymptotic*

$$\lim_{x\to\infty} e^{bx}(1 - g(x)) = a \tag{2.37}$$

then the limit

$$\lim_{t\to\infty} u(t, x + ct) = w_c(x)$$

exists and w_c is a travelling wave solution with speed $c = 1/b + \frac{1}{2}b$.

In his article McKean also briefly mentioned a martingale which (adapted for our different notation) we can write as

$$Z_\lambda(t) := \sum_{k=1}^{N(t)} e^{\lambda X_k(t) - (\frac{1}{2}\lambda^2 + 1)t}, \tag{2.38}$$

and gave a brief proof that this can be used to define the travelling waves w_c, for $c := 1/\lambda + \frac{1}{2}\lambda$, via

$$w_c(x) = \mathbb{E}^0\big(e^{-e^{\lambda x}Z(\infty)}\big).$$

2.6.3 FKPP systems for the finite-type model

For the finite-type model from earlier sections, with types $I = \{1, \ldots, n\}$, we arrive at a FKPP *system* of equations:

$$\frac{\partial f}{\partial t} = \frac{1}{2} A \frac{\partial^2 f}{\partial x^2} + R(f^2 - f) + \theta Q f, \tag{2.39}$$

where f is a vector-valued function from $[0, \infty) \times \mathbb{R}$ to $\mathbb{R}^n$, and A and R are the $n \times n$ diagonal matrices $A := \mathrm{diag}[a(1), \ldots, a(n)]$ and $R := \mathrm{diag}[r(1), \ldots, r(n)]$, as they were in Chapter 2.

The McKean representation takes a modified form:

Theorem 2.6.3 (The McKean representation) *If u satisfies the system (2.39) with $0 \leq u \leq 1$ on $[0, \infty) \times \mathbb{R} \times I$ and with initial condition*

$$u(0, x, y) = f(x, y)$$

then u has a McKean representation*:*

$$u(t, x, y) = \mathbb{E}^{x,y} \prod_{k=1}^{N(t)} f(X_k(t), Y_k(t)).$$

A proof of this in the case $n = 2$ can be found in Champneys *et al* [4]

Using this representation, we later give an explicit (probabilistic) characterization of solutions with Heaviside initial conditions, and also prove that a certain class of solutions of (2.39) converge to travelling waves.

2.6.4 Martingales and Travelling Waves

A solution to (2.39) of the form $f(t, x) = w(x - ct)$, where $w : \mathbb{R} \to \mathbb{R}^n$, is said to be a *travelling wave of speed* c, in which case it must be that

$$\frac{1}{2} A w'' + c w' + R(w^2 - w) + \theta Q w = 0. \tag{2.40}$$

A subset of these are the *monotone travelling waves of speed c that connect 0 to 1*, which along with being a solution to (2.40) have $w(x, \cdot) \to 0$ as $x \to -\infty$, $w(x, \cdot) \to 1$ as $x \to \infty$, and $w'(x, \cdot) > 0$ for all $x \in \mathbb{R}$.

Note for typographical clarity: w is defined as an $\mathbb{R}^n$-valued function of a single, real variable x, and we use $w(x, y)$ to we mean the yth component of $w(x)$.

The following theorem, a corollary of Theorem 2.2.2, shows the close connection between martingales for the branching diffusion and travelling waves for the FKPP equation.

Theorem 2.6.4 *1. For any C^2 function on $\mathbb{R} \times I$, the sum*

$$\sum_{k=1}^{N(t)} h(X_k(t) + ct, Y_k(t)), \qquad (c \in \mathbb{R}) \tag{2.41}$$

is a local martingale if and only if h satisfies the so-called linearized travelling wave system:

$$\frac{1}{2} A h'' + c h' + R h + \theta Q h = 0. \tag{2.42}$$

2. Similarly, if w is a C^2 function on $\mathbb{R} \times I$, then the product

$$\prod_{k=1}^{N(t)} w(X_k(t) + ct, Y_k(t)), \qquad (c \in \mathbb{R}) \tag{2.43}$$

is a local martingale if and only if w satisfies the travelling wave system:

$$\frac{1}{2}Aw'' + cw' + R(w^2 - w) + \theta Qw = 0. \tag{2.44}$$

2.6.5 Results on Monotone Travelling Waves

Existence of MTWs

A proof of the following theorem for the case $n = 2$ can be found in Champneys *et al* [4] – it generalizes without any problems.

Theorem 2.6.5 *For $\lambda \in (\tilde{\lambda}(\theta), 0)$, if we define*

$$w_\lambda(x, y) := \mathbb{E}^{x,y} \exp(-Z_\lambda(\infty))$$

then w_λ is a monotone travelling wave of speed c_λ that connects 0 to 1.

Because of the restriction imposed on the choice of λ, this particular construction cannot yield monotone travelling waves of speeds less than (or equal to) $\tilde{c}(\theta)$. Actually there is an alternative construction that deals with waves of speed $\tilde{c}(\theta)$, but since this requires different techniques with 'derivative' martingales we do not cover these cases and refer the reader to Harris [23] or Kyprianou [35] for examples in the context of BBM.

On the other hand, it is known that *monotone* travelling waves cannot exist at speeds strictly less than $\tilde{c}(\theta)$. Therefore, *with the proviso that here we are going to deal only with travelling waves of speeds strictly greater than $\tilde{c}(\theta)$*, in the next section we prove that the above construction yields the *unique* (up to translation) example of a monotone travelling wave. Note that the proof depends importantly on a result on the asymptotic behaviour of monotone travelling waves proven in section 2.6.5, which is the generalization of the main result in Harris [23].

Uniqueness of MTWs

The following proof is based closely on those given by Champneys et al [4] and by Harris [23], but is included here for completeness and to show how the asymptotic result of the next section (which has not been previously proved in this form) fits into the picture.

Theorem 2.6.6 *The $w_\lambda(x, y)$ defined in Theorem 2.6.5 for $\lambda \in (\tilde{\lambda}(\theta), 0)$ is the unique monotone travelling wave of speed c_λ, up to translation.*

Proof: If we have a monotone travelling wave $w(x, y)$ of speed $c > \tilde{c}(\theta)$, then by Theorem 2.6.4 we deduce that

$$M(t) = \prod_{k=1}^{N(t)} w(X_k(t) + ct, Y_k(t)),$$

defines a local martingale, which being bounded in $[0,1]$ is actually a true martingale. Being positive, it a.s. converges to $M(\infty)$, and furthermore since it is bounded we know that it is UI – so $M(t) = \mathbb{E}(M(\infty)|\mathcal{F}_t)$ and therefore importantly

$$w(x,y) = \mathbb{E}^{x,y} M(\infty).$$

The asymptotic result from the next section states that for any such $w(x,y)$, there is some $\lambda \in (\tilde{\lambda}(\theta), 0)$ and some $\tilde{x} \in \mathbb{R}$ such that $c_\lambda = c$ and

$$1 - w(x,y) \sim v_\lambda(y)\, e^{\lambda(x+\tilde{x})} \qquad \text{as } x \to \infty.$$

Bearing in mind that $-\ln w(x,y) \sim 1 - w(x,y)$ as $x \to \infty$, this means that for any $\varepsilon > 0$ there exists a $D \in \mathbb{R}$ such that

$$(1-\varepsilon) v_\lambda(y) e^{\lambda(x+\tilde{x})} \le -\ln w(x,y) \le (1+\varepsilon) v_\lambda(y) e^{\lambda(x+\tilde{x})}, \qquad \text{for all } x \ge D.$$

Consequently, when *all* the particles of the branching diffusion satisfy: $X_k(t) + c_\lambda t > D, \quad k = 1, \ldots, N(t)$, we will have:

$$(1-\varepsilon) e^{\lambda\tilde{x}} Z_\lambda(t) \le -\ln M(t) = -\sum_{k=1}^{N(t)} \ln w(X_k(t) + ct, Y_k(t)) \le (1+\varepsilon) e^{\lambda\tilde{x}} Z_\lambda(t). \qquad (2.45)$$

This multiple condition on the branching particles is equivalent to a single condition: $L(t) + c_\lambda t > D$, where

$$L(t) := \inf_{k \le N(t)} X_k(t),$$

is the position of the *leftmost* particle in the branching diffusion. The fact that

$$L(t) + c_\lambda t \to +\infty \qquad \text{a.s.} \qquad (2.46)$$

is a direct consequence of the a.s. convergence to 0 of all the martingales Z_β with $\beta \in (-\infty, \tilde{\lambda}(\theta))$ – the inequalities and limit

$$0 \le \exp\{\beta[L(t) + c_\beta t]\} \le \max_{y \in I}\{v_\beta(y)^{-1}\} Z_\beta \to 0$$

force $L(t) + c_\beta t \to \infty$, for each $\beta \in (-\infty, \tilde{\lambda}(\theta))$. Since we can also choose β so that $c_\lambda > c_\beta$, (2.46) is true.

Hence on letting $t \to \infty$ in (2.45), we have

$$(1-\varepsilon) e^{\lambda\tilde{x}} Z_\lambda(\infty) \le -\ln M(\infty) \le (1+\varepsilon) e^{\lambda\tilde{x}} Z_\lambda(\infty),$$

and we deduce that $M(\infty) = \exp(-e^{\lambda\tilde{x}Z_\lambda(\infty)})$ almost surely (since ε was arbitrary). This gives

$$w(x,y) = \mathbb{E}^{x,y} M(\infty) = \mathbb{E}^{x+\tilde{x},y} e^{-Z_\lambda(\infty)} = w_\lambda(x+\tilde{x}, y),$$

which is to say that the travelling wave w is a translate of the travelling wave w_λ from the previous section. □

The Asymptotics of MTWs

As mentioned, Harris [23] first produced an entirely probabilistic proof of the asymptotic shape of the wavefront of travelling waves for the FKPP equation. Our proof generalizes his analysis to the case of finite-type BBM.

Theorem 2.6.7 *Suppose that w is a monotone travelling wave of speed $c > \tilde{c}(\theta)$ that connects 0 to 1. Then there exists $\tilde{x} \in \mathbb{R}$, $\lambda \in (\tilde{\lambda}(\theta), 0)$ such that $c_\lambda = c$ and*

$$1 - w(x,y) \sim v_\lambda(y)\, e^{\lambda(x+\tilde{x})} \qquad \text{as } x \to \infty, \qquad y = 1, \dots, n.$$

Proof: First of all, it is clear from Theorem 2.4.11 that there is a unique $\lambda \in (\tilde{\lambda}(\theta), 0)$ such that $c_\lambda = c$.

The proof is based on the single-particle model. If we write $u(\cdot, y) = 1 - w(\cdot, y)$, then the equation satisfied by u is:

$$\frac{1}{2}Au'' + c_\lambda u' + R(u - u^2) + \theta Q u = 0. \tag{2.47}$$

On the other hand, the generator of the process $(\xi_t + c_\lambda t, \eta_t)$ is:

$$\Big\{\Big(\mathcal{H} + c_\lambda \frac{\partial}{\partial x}\Big) F\Big\}(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + c_\lambda \frac{\partial F}{\partial x} + \theta \sum_{j \in I} Q(y,j)F(x,j).$$

Thus the Feynman-Kac formula suggests that

$$M_t := u(\xi_t + c_\lambda t, \eta_t) \exp\Big(\int_0^t r(\eta_s)(1 - u(\xi_s + c_\lambda s, \eta_s))\,\mathrm{d}s\Big), \tag{2.48}$$

is a positive $\mathbb{P}$-martingale. This is in fact true. Being positive, this martingale will converge $\mathbb{P}$-a.s.

We now use the change of measure introduced in section 2.3.1 to derive another martingale from M_t in which the integral term is much easier to deal with.

Setting $v(x,y) := v_\lambda(y)^{-1}e^{-\lambda x}u(x,y)$, then because $u(x,y) = \mathbb{E}^{x,y}\{M_t\}$ we can write:

$$\begin{aligned} v(x,y) &= \mathbb{E}^{x,y}\{v_\lambda(y)^{-1}e^{-\lambda x}M_t\} \\ &= \mathbb{E}^{x,y}\{v(\xi_t + c_\lambda t, \eta_t)e^{-\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s} \times \frac{\zeta_\lambda(t)}{\zeta_\lambda(0)}\} \\ &= \mathbb{E}_\lambda^{x,y}\{v(\xi_t + c_\lambda t, \eta_t)e^{-\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s}\}. \end{aligned}$$

from which we deduce that

$$v(\xi_t + c_\lambda t, \eta_t)e^{-\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s} \tag{2.49}$$

is a $\mathbb{P}_\lambda$-martingale, which being positive will ($\mathbb{P}_\lambda$a.s.) converge.

Below we show that the integral term in the above (2.49) is finite, but our proof relies on the process $\xi_t + c_\lambda t$ drifting off to $+\infty$ under $\mathbb{P}_\lambda$. We here state that

$$\lim_{t \to \infty} \frac{1}{t}(\xi_t + c_\lambda t) = -\lambda c'_\lambda, \qquad (\mathbb{P}_\lambda\text{-a.s.})$$

and leave the proof until Corollary 4.3.6 in section 4.3.1. Therefore, since $\lambda \in (\tilde{\lambda}(\theta), 0]$ it follows from the structure of c_λ laid out in Theorem 2.4.11 that under $\mathbb{P}_\lambda$ the diffusion $\xi_t + c_\lambda t$ drifts off to $+\infty$.

Now we show that the integral $\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s$ converges almost surely under $\mathbb{P}_\lambda$. Since a positive martingale must either converge to zero or some positive value, it follows that taking logarithms of (2.49) and dividing by $\xi_t + c_\lambda t$ gives:

$$\limsup_{t\to\infty} \left\{ \frac{\ln v(\xi_t + c_\lambda t, \eta_t)}{\xi_t + c_\lambda t} - \frac{1}{\xi_t + c_\lambda t} \int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s \right\} \le 0, \quad \mathbb{P}_\lambda\text{a.s.}$$

For each $y \in I$, $u(x, y) \to 0$ as $x \to -\infty$; combined with the fact that $\xi_t + c_\lambda t \to +\infty$ almost surely under $\mathbb{P}_\lambda$ we deduce that

$$\frac{1}{t}\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\mathrm{d}s \to 0, \qquad \mathbb{P}_\lambda\text{a.s..}$$

The second term in the lim sup above is just the product of this and the reciprocal of $\frac{1}{t}(\xi_t + c_\lambda t)$, (which we know converges $\mathbb{P}_\lambda$-a.s. to $-\lambda c'(\lambda)$), and so in the limit it is zero and can be removed from the lim sup, leaving:

$$\limsup_t \frac{\ln v(\xi_t + c_\lambda t, \eta_t)}{\xi_t + c_\lambda t} \le 0, \qquad \mathbb{P}_\lambda\text{a.s.}$$

which in terms of $u(\xi_t + c_\lambda t, \eta_t)$ means

$$\limsup_t \frac{\ln u(\xi_t + c_\lambda t, \eta_t)}{\xi_t + c_\lambda t} \le \lambda, \qquad \mathbb{P}_\lambda\text{a.s.}$$

Again using the fact that $\frac{1}{t}(\xi_t + c_\lambda t) \to -\lambda c'(\lambda)$, this becomes:

$$\limsup_t \frac{\ln u(\xi_t + c_\lambda t, \eta_t)}{t} \le -\lambda^2 c'(\lambda), \qquad \mathbb{P}_\lambda\text{a.s.}$$

which, along with the fact that $r(\cdot)$ is bounded on I, is enough to establish that

$$\int_0^\infty r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s < +\infty \qquad \text{a.s. under } \mathbb{P}_\lambda \tag{2.50}$$

Thus the exponential term in (2.49) is finite as $t \to \infty$, demonstrating that the term we are most interested in,

$$v(\xi_t + c_\lambda t, \eta_t)$$

must converge a.s. under $\mathbb{P}_\lambda$. The following lemma is proved in the following section:

Lemma 2.6.8 *For each $(x, y) \in E$ the tail algebra is trivial under $\mathbb{P}_\lambda^{x,y}$, and therefore the above process actually converges to a constant value:*

$$v(\xi_t + c_\lambda t, \eta_t) \to K \qquad \mathbb{P}_\lambda\textit{a.s.},$$

for some $K \ge 0$.

Hence it only now remains to prove that K is non-zero. We have shown that (2.49) is a $\mathbb{P}_\lambda$-martingale. *If* we knew that it was uniformly integrable we could write

$$\begin{aligned} v(x,y) &= \mathbb{E}_\lambda^{x,y}\left\{v(\xi_t + ct, \eta_t)\exp\left(-\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s\right)\right\} \\ &= K\,\mathbb{E}_\lambda^{x,y}\left\{\exp\left(-\int_0^\infty r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s\right)\right\} \end{aligned}$$

and deduce that K is strictly positive, since $v(x,y)$ is.

Lemma 2.6.9 *The $\tilde{\mathbb{Q}}$-martingale*

$$v(\xi_t + c_\lambda t, \eta_t)e^{-\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)\,\mathrm{d}s} \tag{2.51}$$

is uniformly integrable.

Proof of Lemma: We have already shown that the term

$$e^{-\int_0^t r(\eta_s)u(\xi_s + c_\lambda s, \eta_s)}$$

remains bounded as $t \to \infty$, so the UI property of this martingale really depends on the behaviour of $v(\xi_t + c_\lambda t, \eta_t)$.

Given the finite number of types it will follow that (2.51) is UI once we show that $v(\cdot, y)$ is bounded on $\mathbb{R}$ for each $y \in I$. We recall that $v(x,y) := v_\lambda(y)^{-1}e^{-\lambda x}u(x,y)$ and $v(x,y) \to 0$ as $x \to -\infty$ for each $y \in I$. Therefore we only need to show boundedness as $x \to \infty$, and our aim is to use a contradiction proof to obtain

$$v(x,y) \to K \text{ as } x \to \infty, \text{for each } y \in I. \tag{2.52}$$

Therefore we start out assuming (without loss of generality) that $v(x,1)$ does not converge to K as $x \to \infty$.

Let $\varepsilon > 0$ be given; then because $v(x,1)$ does not converge, there must be a sequence of points $\{x_n\}$ such that

$$x_n \to \infty \quad \text{and} \quad |v(x_n,1) - K| > \varepsilon, \forall n. \tag{2.53}$$

As it stands, (2.53) doesn't fit well with a probabilistic argument, so we show that in fact we can 'expand' this to intervals around each x_n. We can write (2.53) as:

$$\text{for each } x_n, \text{ either} \quad v(x_n,1) - K > \varepsilon \quad \text{or} \quad K - v(x_n,1) > \varepsilon$$

and for example suppose that at some particular x_n we in fact have $v(x_n,1) - K > \varepsilon$. Since $u(\cdot,y)$ is a decreasing function

$$\frac{u(x_n,1)}{e^{\lambda x_n}} \le \frac{u(x,1)}{e^{\lambda x}} \times e^{\lambda(x-x_n)} \qquad \text{if } x < x_n, \tag{2.54}$$

whence

$$v(x,1) \ge e^{-\lambda(x-x_n)}v(x_n,1) \ge e^{-\lambda(x-x_n)}(K+\varepsilon) \qquad \text{if } x < x_n. \tag{2.55}$$

Furthermore, since $\lambda < 0$ it is possible to find some $\delta = \delta(\lambda)$ *independently* of x_n such that

$$e^{-\lambda(x-x_n)}(K+\varepsilon) > K + \frac{1}{2}\varepsilon \quad \text{for all } x \in (x_n - \delta, x). \tag{2.56}$$

We could have carried out a similar argument for the case $K - v(x_n, 1) > \varepsilon$, and therefore the contradiction assumption (2.53) is actually equivalent to the more useful:

$$|v(x,1) - K| > \varepsilon, \quad \text{whenever } x \in (x_n - \delta, x_n + \delta). \tag{2.57}$$

Intuitively this conflicts with the probability: we know that $v(\xi_t + c_\lambda t, \eta_t)$ converges to K as $t \to \infty$ ($\mathbb{P}_\lambda$-a.s.), but (2.57) implies that after a finite time (with probability 1) the type process η_t must never again jump to type 1 whilst the diffusion $\xi_t + c_\lambda t$ is in any of the remaining intervals $(x_n - \delta, x_n + \delta)$. We make this precise with a Borel-Cantelli-type argument.

There is always a non-zero probability (w.r.t. $\mathbb{P}_\lambda$) that the type of the process will change between the time the diffusion part $\xi_t + c_\lambda t$ hits x_n and when it leaves the interval $(x_n - \delta, x_n + \delta)$ – both events being guaranteed (for large enough n) because under $\mathbb{P}_\lambda$ the process $\xi_t + c_\lambda t$ drifts to $+\infty$. This probability does not depend on n because the process $\xi_t + c_\lambda t$ is space-homogeneous, and furthermore since the state space I of η_t is finite we can define

$$p := \min_{y \in I} \mathbb{P}_\lambda(\eta_t \text{ changes state before } \xi_t + c_\lambda t \text{ leaves } (x_n - \delta, x_n + \delta) | \eta_0 = y, \xi_0 = x_n)$$

and know that $p > 0$. If we likewise define a sequence of *independent* events, for each $n \in \mathbb{N}$,

$$E_n := \{\eta_t \text{ changes state between } \xi_t + ct \text{ hitting } x_n \text{ and leaving } (x_n - \delta, x_n + \delta)\}$$

then it is clear that $\mathbb{P}_\lambda(E_n) \geq p$, and an application of the second Borel-Cantelli lemma gives

$$\mathbb{P}_\lambda(E_n, i.o.) = 1.$$

This contradicts the assumption (2.57), whence

$$v(x,y) \to K, \text{ as } x \to \infty \qquad \text{for each } y \in I.$$

Therefore $v(x,y)$ is bounded on $\mathbb{R} \times I$, and so (2.51) is uniformly integrable.

End of Lemma.

Thus with the UI property proven, we finish by setting $\tilde{x} = \lambda^{-1} \log K$, leaving the required result:

$$\lim_{x \to \infty} \frac{u(x,y)}{v_\lambda(y) e^{\lambda(x+\tilde{x})}} = 1, \qquad y \in \{1, \ldots, n\}.$$

□

A coupling proof of Lemma 2.6.8

We define the following σ-algebras:

$$\mathcal{F}'_t := \sigma((\xi_s s, \eta_s s) : s \geq t),$$

$$\mathcal{T} := \textstyle\bigcap_{t \geq 0} \mathcal{F}'_t.$$

and call $\mathcal{T}$ the *tail σ-algebra.* We show that

$$\lim_{t\to\infty} v(\xi_t + c_\lambda t, \eta_t),$$

(which is clearly $\mathcal{T}$-measurable) is actually ($\mathbb{P}_\lambda$-a.s.) a constant by showing that $\mathcal{T}$ is trivial under each $\mathbb{P}_\lambda^{(x,y)}$ – the measure under which the process (ξ_t, η_t) starts in some state (x, y).

The following argument comes from Rogers [50]. Suppose that Z is some bounded, $\mathcal{T}$-measurable random variable. Then

$$Z = \lim_{t\to\infty} \mathbb{E}_\lambda^{x,y}(Z|\mathcal{F}_t).$$

We are going to show that the conditional expectations $\mathbb{E}_\lambda^{x,y}(Z|\mathcal{F}_t)$ are constant, so that the above limit will force us to deduce that Z is a constant RV – which is to say that $\mathcal{T}$ is trivial under $\mathbb{P}_\lambda^{x,y}$.

First of all, because Z is also in $L^\infty(\mathcal{F}'_t)$, it follows that this conditional expectation actually depends only on the present value of the process (ξ_t, η_t):

$$\mathbb{E}_\lambda^{x,y}(Z|\mathcal{F}_t) = \mathbb{E}_\lambda^{x,y}(Z|(\xi_t, \eta_t)).$$

This right-hand side defines a collection of bounded functions on E: if $z \in E$ then

$$f_t(z) := \mathbb{E}_\lambda^{x,y}(Z|(\xi_t, \eta_t) = z), \qquad \text{for } t \in [0, \infty).$$

An application of the tower property and the Markov property gives the important property of such f_t (which Rogers calls *tail functions* in [50]):

$$f_t(z) = \mathbb{E}_\lambda^{x,y}\big(\mathbb{E}_\lambda^{x,y}(Z|(\xi_{t+s}, \eta_{t+s}))|(\xi_t, \eta_t) = z\big) = P_s f_{t+s}(z)$$

where P_s is the transition semigroup of our Markov process, with, for $z \in E$

$$P_s f_{t+s}(z) := \int_E f_{t+s}(w) P_s(z, \mathrm{d}w),$$

so that $P_s(z, \cdot)$ is the corresponding transition function on E.

We can therefore write, for z and z' in E:

$$\begin{aligned} \big|f_t(z) - f_t(z')\big| &= \big|P_s f_{t+s}(z) - P_s f_{t+s}(z')\big| \\ &= \Big|\int_E f_{t+s}(w) P_s(z, \mathrm{d}w) - \int_E f_{t+s}(w) P_s(z', \mathrm{d}w)\Big| \\ &\leq \|P_s(z, \cdot) - P_s(z', \cdot)\| \ \|f_{t+s}\|_\infty \end{aligned}$$

Here the norm refers to the total variation norm for bounded, signed measures on the state space E. Thus if we can show that $\|P_s(z, \cdot) - P_s(z', \cdot)\| \to 0$ as $s \to \infty$ then it will follow that the functions $f_t(\cdot)$ are constant on E – implying that the conditional expectations are constant on E.

Since our process (ξ_t, η_t) is made up from an ergodic Markov chain and a one-dimensional diffusion, it is natural to use a coupling argument to estimate $\|P_t(z, \cdot) - P_t(z', \cdot)\|$. The idea is to set up two copies[1] of the single-particle model (ξ_t, η_t) and (ξ_t'', η_t'') on a single measure space $(\hat{\Omega}, \mathcal{F}_t, \mathbb{P})$ such that one process starts at z and the other at z' and they meet and stick together within some finite time: ie. for some random time S which is $\mathbb{P}$-a.s. finite we have:

$$(\xi_t, \eta_t) = (\xi_t'', \eta_t'') \quad \text{for } t \geq S, \tag{2.58}$$

The standard coupling inequality then tells us that

$$\|P_t(z, \cdot) - P_t(z', \cdot)\| \leq 2\,\mathbb{P}(S > t),$$

and from our construction we will have *forced* $\mathbb{P}(S > t) \to 0$ as $t \to \infty$, and so get our result.

Now we build the required processes. So let $z = (u, v)$ and $z' = (u', v')$ be two different starting points in our state space $E = \mathbb{R} \times I$.

Let $(\hat{\Omega}, \mathcal{F}_t, \mathbb{P})$ be a measure space on which we have four independent processes:

- $\eta(t)$, an MC on I with Q-matrix Q_λ, starting in state v.
- $\eta'(t)$, an MC on I with Q-matrix Q_λ, starting in state v'.
- $B_1(t)$ a Brownian motion on $\mathbb{R}$ starting at u.
- $B_2(t)$ a Brownian motion on $\mathbb{R}$ starting at u'.

Note that we here insist on independence because we *know* that independent BMs and independent ergodic MCs couple and we can use this to our advantage.

The first diffusion ξ_t is quite easily constructed by time-scaling one of the brownian motions:

$$\xi_t := B_1\left(\int_0^t a(\eta_s)\,\mathrm{d}s\right),$$

and we have our first process (ξ_t, η_t).

We build the second process (ξ_t'', η_t'') in stages. First we couple the types by introducing the following MC on I:

$$\eta_t'' := \begin{cases} \eta_t' & \text{if } t < T \\ \eta_t & \text{if } T \leq t \end{cases}$$

where $T := \inf\{t \geq 0 : \eta_t' = \eta_t\}$. Note that $\mathbb{P}_\lambda(T < \infty) = 1$. The Strong Markov Property guarantees that η_t'' is an MC on I with Q-matrix Q_λ. We now build an 'intermediate' diffusion ξ_t' that we shall cut-and-paste in a moment to get our desired diffusion ξ_t'':

$$\xi_t' := B_2\left(\int_0^t a(\eta_s'')\,\mathrm{d}s\right).$$

Because the *types* of the two diffusions ξ_t and ξ_t' are *identical* after time T, we can write: if $t > T$ then

$$\begin{aligned} \xi_t &= \xi_T + \hat{B}_1\left(\textstyle\int_T^t a(\eta_s)\,\mathrm{d}s\right), \\ \xi_t' &= \xi_T' + \hat{B}_2\left(\textstyle\int_T^t a(\eta_s)\,\mathrm{d}s\right) \end{aligned} \tag{2.59}$$

[1] Note that it is not necessary for the two copies to be independent – see Lindvall [38] for details.

where both $\hat{B}_1$ and $\hat{B}_2$ are independent brownian motions on $\mathbb{R}$ defined for example as:

$$\hat{B}_1(r) := B_1\left(\int_0^T a(\eta_s)\,\mathrm{d}s + r\right) - B_1\left(\int_0^T a(\eta_s)\,\mathrm{d}s\right).$$

We know that two independent brownian motions in $\mathbb{R}$ started at different points ξ_T and ξ'_T will $\mathbb{P}_\lambda$-a.s. meet within a finite amount of time (see Rogers and Williams [52] or Rogers [50]). It is therefore clear that ξ_t and ξ'_t will also meet because the same time-scaling is applied to both $\hat{B}_1$ and $\hat{B}_2$ in (2.59); that this happens within a finite amount of time after T is guaranteed by the ergodic fact that ($\mathbb{P}_\lambda$-a.s.):

$$\int_T^t a(\eta_s)\,\mathrm{d}s \to \infty, \quad \text{as } t \to \infty.$$

Therefore we can finally define our diffusion:

$$\xi''_t := \begin{cases} \xi'_t & \text{if } t < S \\ \xi_t & \text{if } S \leq t \end{cases}$$

where $S := \inf\{t > T : \xi'_t = \xi_t\}$ is $\mathbb{P}_\lambda$-a.s. finite. Another application of the Strong Markov Property guarantees that $t \mapsto \xi''_t$ is the right process, and our proof is complete – we have constructed the two processes required at (2.58).

□

Part II

A new spine approach to $\mathcal{L}^p$ convergence

Chapter 3

A new formulation of the spine approach

In chapter 4 we are going to give spine proofs of the $\mathcal{L}^p$-convergence Theorem 2.5.1 that we just finished proving in the previous chapter; in fact we shall there carry out spine proofs of similar theorems for a total of three different models of branching diffusions. But before embarking on these proofs, in this chapter we present a new formalization of the spine approach that improves the scheme originally laid out by Lyons *et al* [41, 40, 34], and used also in Kyprianou [35] and Olofsson [47], to mention some more recent spine-based studies.

In the first instance our alternative formulation differs from the Lyons *et al* scheme where one of the measures they defined did not have finite mass and could therefore not be normalized to be a probability measure; in our formulation all measures are probability measures and therefore measure changes are carried out by *martingales.* The new relationships that follow from these martingales are crucial in obtaining the large-deviations results in the final part of this thesis.

Our earlier work, covered in chapter 2, was based on the single-particle ideas used by Harris and Williams [21] and also in Champneys *et al* [4]. There we used the martingales of the single-particle model and the whole branching model:

$$\begin{aligned}
\zeta_\lambda(t) &:= e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_s) e^{\lambda\xi_t - E_\lambda t}, \\
Z_\lambda(t) &:= \sum_{u \in N_t} v_\lambda(Y_u(t)) e^{\lambda X_u(t) - E_\lambda t},
\end{aligned}$$

that were related to each other via the Many-to-One theorem. Our initial workings with spine ideas laid out in the Lyons *et al* papers suggested a closer relationship between these two and in fact have led us to introduce a more substantial change to the Lyons *et al* scheme through our use of sub-filtrations on the underlying space of sample trees with spines in contrast to their approach of *marginalizing.* In this way, with the sub-filtrations giving a formal basis, we show how the single-particle model can be incorporated into the branching particles through the spine, with the result that the relationship between the whole branching collection and the single-particle model is represented through the conditional-expectation operation.

This new point of view gives substantial improvements to both the Harris and Williams single-particle approach and the Lyons *et al* spine approach, since it combines them into a more powerful object which can be interpreted from either point of view. In the first instance this gives us the correct formal basis in which to express the relationship between the single-particle martingale ζ_λ and the additive martingale Z_λ that we used in Chapter 2. This same idea will work in more generality and becomes a key element to developing *new martingales* for branching diffusions that can be very powerful in proving large-deviations results, as we explore further in Part III.

A new and interesting aspect of our formulation is the relation that becomes clear between the spine and the 'Gibbs-Boltzmann' weightings for the branching particles. Such weightings are well known in the theory of branching process, and Harris [24] contains some analysis for a model of a typed branching diffusion that we shall be looking at in later chapters. Here we explain how these weightings can be interpreted as a conditional expectation of a spine event, and then develop this further to show how such conditional expectations can obtain a new and very useful interpretation of the additive operations previously seen only within the context of the Kesten-Stigum theorem and related problems.

Furthermore, as a consequence of the Gibbs-Boltzmann weightings we obtain a substantially easier proof of an improved version of the Many-to-One theorem used by Harris and Williams.

Overview of spine ideas

One of the central elements of the spine approach is to interpret the behaviour of a branching process under a new measure defined in terms of the additive martingales. Such an interpretation was first laid out by Chauvin and Rouault [8] in the case of branching Brownian motion, and we briefly review the main ideas on a heuristic level.

Consider a branching Brownian motion (BBM) with constant branching rate r, which is the branching process whereby particles diffuse independently according to a Brownian motion and at any moment undergo fission at a rate r to produce two particles. We suppose that the probabilities of this are $\{P^x : x \in \mathbb{R}\}$ so that P^x is a measure defined on the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ such that it is the law of the process initiated from a single particle positioned at x. Suppose that the configuration of this branching Brownian motion at time t is given by the $\mathbb{R}$-valued point process $\mathbb{X}_t := \{X_u(t) : u \in N_t\}$ where N_t is the set of individuals alive at time t. It is well known that for any $\lambda \in \mathbb{R}$,

$$Z_\lambda(t) := \sum_{u\in N_t} e^{-rt} e^{\lambda X_u(t) - \frac{1}{2}\lambda^2 t} \tag{3.1}$$

defines a *strictly-positive* P-martingale, so $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ is almost surely finite under P^x. The important contribution of Chauvin and Rouault [8] was to determine a pathwise construction of the measure $\mathbb{Q}^x_\lambda$ such that

$$\left.\frac{\mathrm{d}\mathbb{Q}^x_\lambda}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)}, \tag{3.2}$$

where the term $Z_\lambda(0)$ acts as a normalizing factor.

Theorem 3.0.10 *The measure $\mathbb{Q}^x_\lambda$ defined at (3.2) is equivalent to the following pathwise construction:*

- *starting from position x, the original ancestor diffuses according to a Brownian motion on $\mathbb{R}$ with drift λ;*
- *at rate $2r$ the particle undergoes fission producing two particles;*
- *with equal probability, one of these two particles is selected;*
- *this chosen particle repeats stochastically the behaviour of the parent;*
- *the other particle initiates, from its birth position, an independent copy of a $P^\cdot$ branching Brownian motion with branching rate r.*

We briefly note that for the third point the concept of randomly choosing between two particles must imply a *difference* between the two particles. In our formal approach to the underlying probability spaces each particle will be labelled according to the Ulam-Harris scheme, and therefore the idea of choosing a particle is really a question of choosing between *labels.*

The chosen line of descent in such pathwise constructions of the measure, here $\mathbb{Q}_\lambda$, has come to be known as the *spine* because it can simply be thought of as the backbone of the branching process $\mathbb{X}_t$ from which all particles are born. Although Chauvin and Rouault's work on the measure change continued in a paper co-authored with Wakolbinger [9], where the new measure is interpreted as the result of building a conditioned tree using the concepts of Palm measures, it wasn't until the so-called 'conceptual proofs' of Lyons, Kurtz, Peres and Pemantle published around 1995 ([41, 40, 34]) that the spine approach really began to crystalize. These papers laid out a formal basis for spines using a series of new measures on two underlying spaces of sample trees with and without distinguished lines of descent (the spine). Of particular interest to this thesis is the paper by Lyons [40] which gave a spine proof of the $\mathcal{L}^1$-convergence of an additive martingale for a branching random walk. Here the idea of using the martingale as a measure-changing Radon-Nikodym derivative was brought together with a previously known measure-theoretic result that allows us to deduce the behaviour of the change-of-measure martingale in the original measure by investigating its behaviour in the second measure, and the new *spine decomposition* of the martingale was first to be a key to using the intuition provided by Chauvin and Rouault's pathwise construction of the new measure.

Similar ideas have more recently been used by Kyprianou [35] to further investigate the $\mathcal{L}^1$-convergence of the BBM martingale (3.1), and we discuss these more fully in the next chapter.

In this thesis we shall be using spines in a number of different branching diffusions, and therefore we base the presentation of our new formulation on a non-specific Markov branching model which is more than general enough to cover all our cases. Here, particles move independently in a space J as a stochastic copy of some given stochastic process Ξ_t, and at a location-dependent rate undergo fission to produce a location-dependent random number of offspring that each carry on this branching behaviour independently.

Definition 3.0.11 (A general branching Markov process) *We suppose that three initial elements are given to us:*

- *a Markov process Ξ_t in a measurable set $(J, \mathcal{B})$,*
- *a measurable function $R : J \to [0, \infty)$,*
- *for each $x \in J$ we are given a random variable $A(x)$ whose probability distribution on the natural numbers $\{0, 1, \ldots\}$ is $P\big(A(x) = k\big) = p_k(x)$, and whose mean is $m(x) := \sum_{k=0}^{\infty} kp_k(x)$.*

From these ingredients we can build a branching process in J according to the following recipe:

- *Each particle of the branching process will live, move and die in this space $(J, \mathcal{B})$, and if an individual u is alive at time t we refer to its location in J as $X_u(t)$. Therefore the time-t configuration of the branching process is a J-valued point process $\mathbb{X}_t := \{X_u(t) : u \in N_t\}$ where N_t denotes the collection of all particles alive at time t.*
- *For each individual u, the stochastic behaviour of its motion in J is an independent copy of the given process Ξ_t.*
- *The function $R : J \to [0, \infty)$ determines the rate at which each particle dies: given that u is alive at time t, its probability of dying in the interval $[t, t + \mathrm{d}t)$ is $R(X_u(t))\mathrm{d}t + o(\mathrm{d}t)$.*
- *If a particle u dies at location $x \in J$ it is replaced by $1 + A_u$ particles all positioned at x, where A_u is an independent copy of the random variable $A(x)$. All particles, once born, progress independently of each other.*

We suppose that the probabilities of this branching process are $\{P^x : x \in J\}$ so that under P^x one initial ancestor starts out at x.

We shall first give a formal construction of the underlying probability space, made up of the sample trees of the branching process $\mathbb{X}_t$ in which the spines are the distinguished lines of descent. Once built, this space will be filtered in a natural way by the underlying family relationships of each sample tree, the diffusing branching particles and the diffusing spine, and then in section 3.2 we shall explain how we can define new probability measures $\tilde{P}^x$ that extend each P^x up to the finest filtration that contains all information about the spine and the branching particles.

Since the presentation in terms of this general model can sometimes make the results look more difficult than they actually are, we tend to use the familiar example of the finite-type branching diffusion from chapter 2 to first introduce the ideas, then follow up with the general formulation. For example, to deal with this finite-type model we would take the process Ξ_t to be the single-particle process (ξ_t, η_t) described at 2.3.1, which lives in the space $J = \mathbb{R} \times I$. The birth rate of that model was given by the function $R(y)$ for all $y \in I$, and since we only dealt with binary branching we would suppose that $P(A(y) = 1) = 1$ for all $y \in I$.

Finally, we remark that much of the *notation* that we use for the underlying space of trees, the filtrations and the measures is closely based on that used in Kyprianou [35].

3.1 The underlying space for spines

3.1.1 Marked Galton-Watson trees with spines

The set of Ulam-Harris labels is to be equated with the set Ω of finite sequences of strictly-positive integers:

$$\Omega := \{\emptyset\} \cup \bigcup_{n \in \mathbb{N}} (\mathbb{N})^n,$$

where we take $\mathbb{N} = \{1, 2, \ldots\}$. For two words $u, v \in \Omega$, uv denotes the concatenated word ($u\emptyset = \emptyset u = u$), and therefore Ω contains elements like '213' (or '$\emptyset$213'), which we read as 'the individual being the 3rd child of the 1st child of the 2nd child of the initial ancestor $\emptyset$'. For two labels $v, u \in \Omega$ the notation $v < u$ means that v is an *ancestor* of u, and $|u|$ denotes the length of u. The set of all ancestors of u is equally given by

$$\{v : v < u\} = \{v : \exists w \in \Omega \text{ such that } vw = u\}.$$

Collections of labels, ie. subsets of Ω, will therefore be groups of individuals. In particular, a subset $\tau \subset \Omega$ will be called a *tree* if:

1. $\emptyset \in \tau$,
2. if $u, v \in \Omega$, then $uv \in \tau$ implies $u \in \tau$,
3. for all $u \in \tau$, there exists $A_u \in 0, 1, 2, \ldots$ such that $uj \in \tau$ if and only if $1 \leq j \leq 1 + A_u$, (where $j \in \mathbb{N}$).

That is just to say that a tree:

1. has a single initial ancestor $\emptyset$,
2. contains all ancestors of any of its individuals v,
3. has the $1 + A_u$ children of an individual u labelled in a consecutive way,

and is therefore just what we imagine by the picture of a family tree descending from a single ancestor. Note that the '$1 \leq j \leq 1 + A_u$' condition in 3 means that each individual has *at least* one child, so that in our model we are insisting that trees *never die out.*

The set of all trees will be called $\mathbb{T}$. Typically we use the name τ for a particular tree, and whenever possible we will use the letters u or v or w to refer to the labels in τ, which we may also refer to as *nodes of* τ or *individuals in* τ or just as *particles.*

Each individual should have a *location* in J at each moment of its *lifetime*. Since a Galton-Watson tree $\tau \in \mathbb{T}$ in itself can express only the *family* structure of the individuals in our branching model, in order to give them these extra features we suppose that each individual $u \in \tau$ has a mark (X_u, σ_u) associated with it which we read as:

- $\sigma_u \in \mathbb{R}^+$ is the *lifetime* of u, which determines the *fission time* of particle u as $S_u := \sum_{v \leq u} \sigma_v$ (with $S_\emptyset := \sigma_\emptyset$). The times S_u may also be referred to as the *death* times;
- $X_u : [S_u - \sigma_u, S_u) \to J$ gives the *location* of u at time $t \in [S_u - \sigma_u, S_u)$.

To avoid ambiguity, it is always necessary to decide whether a particle is in existence or not at its death time.

Remark 3.1.1 *Our convention throughout will be that a particle u dies 'just before' its death time S_u (which explains why we have defined $X_u : [S_u - \sigma_u, S_u) \to \cdot$ for example). Thus at the time S_u the particle u has* disappeared, *replaced by its $1 + A_u$ children which are all alive and ready to go.*

We denote a single marked tree by (τ, X, σ) or (τ, M) for shorthand, and the set of all marked Galton-Watson trees by $\mathcal{T}$:

- $\mathcal{T} := \Big\{(\tau, X, \sigma) : \tau \in \mathbb{T}$ and for each $u \in \tau, \sigma_u \in \mathbb{R}^+, X_u : [S_u - \sigma_u, S_u) \to J\Big\}$.
- For each $(\tau, X, \sigma) \in \mathcal{T}$, the set of particles that are alive at time t is defined as $N_t := \{u \in \tau : S_u - \sigma_u \leq t < S_u\}$.

Where we want to highlight the fact that these values depend on the underlying marked tree we write e.g. $N_t((\tau, X, \sigma))$ or $S_u((\tau, M))$.

Any particle $u \in \tau$ that comes into existence creates a *subtree* made up from the collection of particles (and all their marks) that have u as an ancestor – and u is the original ancestor of this subtree.

- $(\tau, X, \sigma)^u_j$, or $(\tau, M)^u_j$ for shorthand, is defined as the *subtree* growing from individual u's jth child uj, where $1 \leq j \leq 1 + A_u$.

This subtree is a marked tree itself, but when considered as a part of the original tree we have to remember that it comes into existence at the space-time location $(X_u(0), S_u - \sigma_u)$ – which is just the space-time location of the death of particle u (and therefore the space-time location of the birth of its child uj).

Before moving on there is a final extension of the notation to be made: for any particle u we extend the definition of X_u from the time interval $[S_u - \sigma_u, S_u)$ to allow all earlier times $t \in [0, S_u)$:

Definition 3.1.2 *Each particle u is alive in the time interval $[S_u - \sigma_u, S_u)$, but we extend the concept of its path in J to all earlier times $t < S_u$:*

$$X_u(t) := \begin{cases} X_u(t) & \text{if } S_u - \sigma_u \leq t < S_u \\ X_v(t) & \text{if } v < u \text{ and } J_v \leq t < S_v \end{cases}$$

Thus particle u inherits the path of its unique line of ancestors, and this simple extension will allow us to later write expressions like $e^{\int_0^t f(s)\,\mathrm{d}X_u(s)}$ whenever $u \in N_t$, without worrying about the birth time of u.

For any given marked tree $(\tau, M) \in \mathcal{T}$ we can identify distinguished lines of descent from the initial ancestor: $\emptyset, u_1, u_2, u_3, \ldots \in \tau$, in which u_3 is a child of u_2, which itself is a child of u_1 which is a child of the original ancestor $\emptyset$. We'll call such a subset of τ a *spine*, and will refer to it as ξ:

- a spine ξ is a subset of nodes $\{\emptyset, u_1, u_2, u_3, \ldots\}$ in the tree τ that make up a unique line of descent. We use ξ_t to refer to the unique node in ξ that that is alive at time t.

In a more formal definition, which can for example be found in the paper by Rouault and Liu [39], a spine is thought of as a point on $\partial\tau$ the boundary of the tree – in fact the boundary is *defined* as the set of all infinite lines of descent. This explains the notation $\xi \in \partial\tau$ in the following definition: we augment the space $\mathcal{T}$ of marked trees to become

- $\tilde{\mathcal{T}} := \Big\{(\tau, M, \xi) : (\tau, M) \in \mathcal{T} \text{ and } \xi \in \partial\tau\Big\}$ is the set of *marked trees with distinguished spines.*

It is natural to speak of the *position of the spine at time t*, which we think of just as the position of the unique node that is in the spine and alive at time t:

- we define the time-t position of the spine as $\xi_t := X_u(t)$, where $u \in \xi \cap N_t$.

By using the notation ξ_t to refer to both the node in the tree and that node's spatial position we are introducing potential ambiguity, but in practice the context will make clear which we intend. However, in case of needing to emphasize, we shall give the node a longer name:

- $\mathrm{node}_t((\tau, M, \xi)) := u$ if $u \in \xi$ is the node in the spine alive at time t,

which may also be written as $\mathrm{node}_t(\xi)$.

Finally, it will later be important to know how many fission times there have been in the spine, or what is the same, to know which generation of the family tree the node ξ_t is in (where the original ancestor $\emptyset$ is considered to be the 0th generation)

Definition 3.1.3 *We define the counting function*

$$n_t = \big|node_t(\xi)\big|,$$

which tells us which generation the spine node is in, or equivalently how many fission times there have been on the spine. For example, if $\xi_t = \big(\emptyset, u_1, u_2\big)$ *then both* $\emptyset$ *and* u_1 *have died and so* $n_t = 2$.

3.1.2 Filtrations

The reader who is already familiar with the Lyons *et al* [34, 40, 41] papers will recall that they used two separate underlying spaces of marked trees *with* and *without* the spines, then marginalized out the spine when wanting to deal only with the branching particles as a whole. Instead, we are going to use just the single underlying space $\tilde{\mathcal{T}}$, but define *four* filtrations of it that encapsulate different knowledge:

- $\mathcal{F}_t$ knows everything that has happened to all the branching particles up to the time t, *but does not know which one is the spine*;
- $\tilde{\mathcal{F}}_t$ knows everything that $\mathcal{F}_t$ knows and also knows which line of descent is the spine (it is in fact the finest filtration);
- $\mathcal{G}_t$ knows only about the spine's motion in J up to time t, but does not actually know which line of descent in the family tree makes up the spine;
- $\tilde{\mathcal{G}}_t$ knows about the spine's motion and also knows which nodes it is composed of. Furthermore it knows about the fission times of these nodes and how many children were born at each time.

Filtration $(\mathcal{F}_t)_{t\geq 0}$

We define a filtration of $\tilde{\mathcal{T}}$ made up of the σ-algebras:

$$\mathcal{F}_t := \sigma\Big((u, X_u, \sigma_u) : S_u \leq t\ ;\ (u, X_u(s) : s \in [S_u - \sigma_u, t]) : t \in [S_u - \sigma_u, S_u)\Big),$$

which in words means that $\mathcal{F}_t$ is generated by all the information regarding the branching particles that have *lived and died* before time t (this is the condition $S_u \leq t$), along with just the information up to time t of those particles *still alive* at time t (this is the $t \in [S_u - \sigma_u, S_u)$ condition). Each of these σ-algebras will be a subset of the limit defined as

$$\mathcal{F}_\infty := \sigma\Big(\bigcup_{t\geq 0} \mathcal{F}_t\Big).$$

Filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$

In order to know about the spine, we make this filtration finer, defining $\tilde{\mathcal{F}}_t$ by adding into $\mathcal{F}_t$ the knowledge of which node is the spine at time t:

$$\tilde{\mathcal{F}}_t := \sigma\big(\mathcal{F}_t, \mathrm{node}_t(\xi)\big), \qquad \tilde{\mathcal{F}}_\infty := \sigma\Big(\bigcup_{t\geq 0} \tilde{\mathcal{F}}_t\Big).$$

Consequently this filtration knows *everything* about the branching process and *everything* about the spine: it knows which nodes make up the spine, when they were born, when they died (ie. the fission times S_u), and their family sizes.

Filtration $(\mathcal{G}_t)_{t\geq 0}$

We define a filtration of $\tilde{\mathcal{T}}$, $\{\mathcal{G}_t\}_{t\geq 0}$, where the σ-algebras

$$\mathcal{G}_t := \sigma(\xi_s : 0 \leq s \leq t), \qquad \mathcal{G}_\infty := \sigma\Big(\bigcup_{t\geq 0} \mathcal{G}_t\Big),$$

are generated by *only* the spatial motion of the spine in the J. Note that the events $G \in \mathcal{G}_t$ do not know which *nodes* of the tree τ actually make up the spine.

Filtration $(\tilde{\mathcal{G}}_t)_{t\geq 0}$

We augment $\mathcal{G}_t$ by adding in information on the nodes that make up the spine (as we did from $\mathcal{F}_t$ to $\tilde{\mathcal{F}}_t$), as well as the knowledge of when the fission times occurred on the spine and how big the families were that were produced:

$$\tilde{\mathcal{G}}_t := \sigma\big(\mathcal{G}_t, (\mathrm{node}_s : s \leq t), (A_u : u < \xi_t)\big), \qquad \tilde{\mathcal{G}}_\infty := \sigma\Big(\bigcup_{t\geq 0} \tilde{\mathcal{G}}_t\Big).$$

Summary

Here is a layout of the relationships between the different filtrations of $\tilde{\mathcal{T}}$:

$$\mathcal{G}_t \subset \tilde{\mathcal{G}}_t \subset \tilde{\mathcal{F}}_t$$
$$\mathcal{F}_t \subset \tilde{\mathcal{F}}_t$$

Importantly, but trivially, we have $\mathcal{G}_t \not\subseteq \mathcal{F}_t$, since the filtration $\mathcal{F}_t$ does not know *which* line of descent makes up the spine.

3.2 Probability measures

Having now carefully defined the underlying space for our probabilities, we remind ourselves of the probability measures:

Definition 3.2.1 *For each $x \in J$, let P^x be the measure on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ such that the filtered probability space $(\tilde{\mathcal{T}}, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0}, P^\cdot)$ is the canonical model for $\mathbb{X}_t$, the branching Markov process described in definition 3.0.11.*

For details of how the measures P^x are formally constructed on the underlying space of trees, we refer the reader to the work of Neveu [43] and Chauvin [7, 5].

All spine approaches rely on building a measure $\tilde{P}^x$ under which the spine is a single genealogical line of descent chosen uniformly from the underlying tree. If we are given a sample tree (τ, M) for the branching process it can be verified that a 'harmonic' choice of which line of descent makes up the spine ξ implies that if $u \in \tau$ then

$$\mathrm{Prob}(u \in \xi) = \prod_{v<u} \frac{1}{1 + A_v}. \tag{3.3}$$

This observation is the key to our method for extending the measures, and for this we make use of the following representation found in Lyons [40].

Theorem 3.2.2 *If f is a $\tilde{\mathcal{F}}_t$-measurable function then we can write:*

$$f = \sum_{u \in N_t} f_u \mathbf{1}_{(\xi_t = u)} \tag{3.4}$$

where f_u is $\mathcal{F}_t$-measurable.

As an example of this, in the case of the finite-typed branching diffusion of chapter 2, such a representation would be:

$$e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t)\, e^{\lambda \xi_t - E_\lambda t} = \sum_{u \in N_t} e^{\int_0^t R(Y_u(s))\,\mathrm{d}s} v_\lambda(Y_u(t))\, e^{\lambda X_u(t) - E_\lambda t} \mathbf{1}_{(\xi_t = u)}. \tag{3.5}$$

Definition 3.2.3 *Given the measure P^x on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty)$ we extend it to the probability measure $\tilde{P}^x$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by defining*

$$\int_{\tilde{\mathcal{T}}} f \;\mathrm{d}\tilde{P}^x := \int_{\mathcal{T}} \sum_{u \in N_t} f_u \prod_{v < u} \frac{1}{1 + A_v} \;\mathrm{d}P^x, \tag{3.6}$$

for each $f \in m\tilde{\mathcal{F}}_t$ with representation like (3.4).

The previous approach to spines, exemplified in Lyons [40], used the idea of *fibres* to get a measure analogous to our $\tilde{P}$ that could measure the spine. In their approach the corresponding measure did not have a finite mass and therefore could not be normalized to become a probability measure like our $\tilde{P}$. Our new idea of using the down-weighting term $1/(1 + A_v)$ in the definition of $\tilde{P}$ (which is just to say that we are using a harmonic measure on the underlying tree) is crucial in ensuring that we do not get an infinite-mass measure, and furthermore implies that in our scheme *all* measure changes are carried out by *martingales*. The importance of having martingales will later be seen in how easy it becomes under our formulation to define new and very powerful martingales for branching diffusions from the standard and well-studied martingales of single-particle diffusions. These new martingales will be crucial in proving the large-deviations result in the final part of this thesis.

Theorem 3.2.4 *This measure $\tilde{P}^x$ really is an extension of P^x in that $P = \tilde{P}|_{\mathcal{F}_\infty}$.*

Proof: If $f \in m\mathcal{F}_t$ then the representation (3.4) is trivial and therefore by definition

$$\int_{\tilde{\mathcal{T}}} f \;\mathrm{d}\tilde{P} = \int_{\tilde{\mathcal{T}}} f \times \Big(\sum_{u \in N_t} \prod_{v < u} \frac{1}{1 + A_v} \Big) \;\mathrm{d}P.$$

However, it can be shown that $\sum_{u \in N_t} \prod_{v<u} \frac{1}{1+A_v} = 1$ by retracing the sum back through the lines of ancestors to the original ancestor $\emptyset$, factoring out the product terms as each generation is passed. Thus

$$\int_{\tilde{\mathcal{T}}} f \,\mathrm{d}\tilde{P} = \int_{\tilde{\mathcal{T}}} f \,\mathrm{d}P.$$

□

Definition 3.2.5 *The filtered probability space* $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty, (\tilde{\mathcal{F}}_t)_{t\geq 0}, \mathbb{P})$ *together with* $(\mathbb{X}_t, \xi_t)$ *will be referred to as the* ***canonical model with spines.***

In chapter 2 we used a separate single-particle model to evaluate some additive expectation calculations. Thus we had assumed the existence of a measure $\mathbb{P}$ and a process (ξ_t, η_t) that behaved stochastically like a 'typical' particle in the typed branching diffusion $\mathbb{X}_t$. In our formalization the *spine* is exactly the single-particle model:

Definition 3.2.6 *We define the measure* $\mathbb{P}$ *on* $(\tilde{\mathcal{T}}, \mathcal{G}_\infty)$ *as the projection of* $\tilde{P}$*:*

$$\mathbb{P}|_{\mathcal{G}_t} := \tilde{P}|_{\mathcal{G}_t}.$$

Under the measure $\mathbb{P}$ the spine process ξ_t has exactly the same law as Ξ_t.

Definition 3.2.7 *The filtered probability space* $(\tilde{\mathcal{T}}, \mathcal{G}_\infty, (\mathcal{G}_t)_{t\geq 0}, \tilde{P})$ *together with the spine process* ξ_t *will be referred to as the* ***single-particle model.***

3.3 Martingales

Starting with the single Markov process Ξ_t that lives in $(J, \mathcal{B})$ we have built $(\mathbb{X}_t, \xi_t)$, a branching Markov process with spines, in which the spine ξ_t behaves stochastically like the given Ξ_t. In this section we are going to show how *any* given martingale $\zeta(t)$ for the spine leads to a corresponding additive martingale for the whole branching model.

We have actually seen an example of this already. In chapter 2 we met two martingales:

$$Z_\lambda(t) := \sum_{u\in N_t} v_\lambda(Y_u(t))e^{\lambda X_u(t) - E_\lambda t}, \tag{3.7}$$

$$\zeta_\lambda(t) := e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t) e^{\lambda\xi_t - E_\lambda t}. \tag{3.8}$$

Just from their very form it has always been clear that they are closely related, and to a degree this was confirmed through the Many-to-One Theorem 2.3.2. What we shall be demonstrating in full generality in this section is that the key to their relationship comes through the following $\tilde{\mathcal{F}}_t$-measurable martingale.

Definition 3.3.1 *We define an* $\tilde{\mathcal{F}}_t$*-measurable martingale:*

$$\tilde{\zeta}_\lambda(t) := \prod_{u<\xi_t}(1+A_u) \times v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t}. \tag{3.9}$$

The important result that we show in this chapter, in a more general form, is that $Z_\lambda(t)$ and $\zeta_\lambda(t)$ are both projections of just conditional expectations of this new martingale $\tilde{\zeta}_\lambda$:

- $Z_\lambda(t) = \tilde{P}(\,\tilde{\zeta}_\lambda(t)\,|\,\mathcal{F}_t)$,
- $\zeta_\lambda(t) = \tilde{P}(\,\tilde{\zeta}_\lambda(t)\,|\,\mathcal{G}_t)$.

We emphasize that this relationship has not previously been formalized, and that it is only *possible* because of our new approach to the definition of $\tilde{P}$ as a *probability* measure, and of our using filtrations to capture the different knowledge generated by the spine and the branching particles.

Furthermore, in the general form that we present below it provides a consistent methodology for using well-known martingales for a single-particle process ξ_t to get new additive martingales for the related branching process. In the third part of this thesis we use these powerful ideas to give substantially easier proofs of large-deviations problems in branching diffusions than have previously been possible.

Suppose that $\zeta(t)$ is a strictly-positive $(\tilde{\mathcal{T}}, (\mathcal{G}_t)_{t\geq 0}, \tilde{P})$-martingale that is the Radon-Nikodym derivative of a new measure $\hat{\mathbb{P}}$ with respect to $\mathbb{P}$; thus it is a $\mathcal{G}_t$-measurable function that is a martingale with respect to the measure $\tilde{P}$. For example in the case of our finite-type branching diffusion this could be the martingale $\zeta_\lambda(t)$ which is $\mathcal{G}_t$-measurable since it refers only to the spine process (ξ_t, η_t).

Definition 3.3.2 *We shall call $\zeta(t)$ the* ***single-particle martingale,*** *since it refers only to the spine ξ (it is $\mathcal{G}_t$-measurable). We suppose that there exists some measure $\hat{\mathbb{P}}$ on $(\tilde{\mathcal{T}}, \mathcal{G}_\infty)$ such that*

$$\left.\frac{\mathrm{d}\hat{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} = \frac{\zeta(t)}{\zeta(0)}.$$

Any $\mathcal{G}_t$-measurable function is immediately an $\tilde{\mathcal{F}}_t$-measurable function and can therefore be used to define an $\mathcal{F}_t$-measurable function via projection according to the representation (3.4), as we explain below. In the case that we use the single-particle martingale $\zeta(t)$, the resulting $\mathcal{F}_t$-measurable function is an additive martingale for the whole branching process via the representation (3.4). We recall that the function $m(x)$ is just the mean number of offspring of the fissions that occur from location x:

$$m(x) := P(A(x)) = \sum_{k=0}^{\infty} kp_k(x).$$

Definition 3.3.3 *Suppose that we can represent the martingale $\zeta(t)$ as*

$$\zeta(t) = \sum_{u\in N_t} \zeta_u(t)\mathbf{1}_{(\xi_t=u)}, \tag{3.10}$$

for $\zeta_u(t) \in m\mathcal{F}_t$, as at (3.4). We can then define an $\mathcal{F}_t$-measurable additive martingale $Z(t)$ as

$$Z(t) := \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t),$$

and refer to $Z(t)$ as the ***branching-particle martingale.***

For clarity we take a moment to discuss this definition of the additive martingale and the terms like $\zeta_u(t)$.

If we return to our familiar martingales (3.7) and (3.8), it is clear that

$$\zeta_\lambda(t) = \sum_{u \in N_t} e^{\int_0^t R(Y_u(s))\,\mathrm{d}s} v_\lambda\big(Y_u(t)\big) e^{\lambda X_u(t) - E_\lambda t} \mathbf{1}_{(\xi_t = u)}. \tag{3.11}$$

The 'ζ_u' terms of (3.10) could be here replaced with a more descriptive notation $\zeta_\lambda[(X_u, Y_u)](t)$, where

$$\zeta_u(t) = \zeta_\lambda[(X_u, Y_u)](t) := e^{\int_0^t R(Y_u(s))\,\mathrm{d}s} v_\lambda\big(Y_u(t)\big) e^{\lambda X_u(t) - E_\lambda t},$$

can be seen to essentially be a functional of the space-type path $(X_u(t), Y_u(t))$ of particle u. In this way the original single-particle martingale ζ_λ is a functional of the space-type path (ξ_t, η_t) of the spine itself and we could write

$$\zeta_\lambda(t) = \zeta_\lambda[(\xi, \eta)](t) = \sum_{u \in N_t} \zeta_\lambda[(X_u, Y_u)](t)\, \mathbf{1}_{(\xi_t = u)}.$$

This is the idea behind the representation (3.10), and in those typical cases where the single-particle martingale is essentially a functional of the paths of the spine ξ_t, as is the case for our $\zeta_\lambda(t)$, we should just think of ζ_u as being that same functional but evaluated over the path $X_u(t)$ of particle u rather than the spine ξ_t. The representation (3.10) is used as a more general way of treating single-particle martingales that perhaps are not such a simple functional of the spine path.

Finally, from (3.11) it is clear that the additive martingale being defined by definition 3.3.3 is our familiar $Z_\lambda(t)$:

$$\begin{aligned} Z_\lambda(t) &= \sum_{u \in N_t} e^{-\int_0^t R(Y_u(s))\,\mathrm{d}s} \zeta_\lambda[(X_u, Y_u)](t) \\ &= \sum_{u \in N_t} v_\lambda\big(Y_u(t)\big) e^{\lambda X_u(t) - E_\lambda t}. \end{aligned}$$

The work of Lyons *et al* [40, 34, 41], that of Chauvin and Rouault [8] and more recently of Kyprianou [35] suggests that when a change of measure is carried out with a branching-diffusion additive martingale like $Z(t)$ it is typical to expect three changes: the spine will gain a drift, its fission rate will be increased and the distribution of its family sizes will be size-biased. In section 3.4.1 we shall confirm this, and now take a separate look at the martingales that could perform these changes, and which we shall combine to obtain a martingale $\tilde{\zeta}(t)$ that will ultimately be used to change the measure $\tilde{P}$.

Theorem 3.3.4 *The expression*

$$\prod_{v < \xi_t} \big(1 + m(\xi_{S_v})\big) e^{-\int_0^t m(\xi_s) R(\xi_s)\,\mathrm{d}s}$$

is a $\tilde{P}$-martingale that will increase the rate at which fissions occur along the spine from $R(\xi_t)$ to $(1 + m(\xi_t))R(\xi_t)$.

Such a change of measure for Poisson processes is discussed in Engländer and Kyprianou [14]. The following type of change of measure has been extensively investigated in the Lyons *et al* papers [41, 40, 34], and was used also in Kyprianou [35] and Olofsson [47], to mention some more recent spine-based studies.

Theorem 3.3.5 *The term* $\prod_{v<\xi_t} \frac{1+A_v}{1+m(\xi_{S_v})}$ *is a* $\tilde{P}$*-martingale that will change the measure by size-biasing the family sizes born from the spine:*

$$\text{if } v < \xi_t, \text{ then} \qquad Prob(A_v = k) = \frac{1+p_k(\xi_{S_v})}{1+m(\xi_{S_v})}.$$

The product of these two martingales with the single-particle martingale $\zeta(t)$ will simultaneously perform the three changes mentioned above.

Definition 3.3.6 *We define a* $\tilde{\mathcal{F}}_t$*-measurable martingale as*

$$\begin{aligned}\tilde{\zeta}(t) &:= \prod_{v<\xi_t} (1+A_v) e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \times \zeta(t) \\ &= \prod_{v<\xi_t} \frac{1+A_v}{1+m(\xi_{S_v})} \times \prod_{v<\xi_t} (1+m(\xi_{S_v}))e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \times \zeta(t). \qquad (3.12)\end{aligned}$$

This martingale is the general form of $\tilde{\zeta}_\lambda(t)$ that we defined earlier at (3.9) for our finite-type model.

The real importance of the size-biasing and fission-rate-increase operations is that they introduce the correct terms into $\tilde{\zeta}(t)$ so that the following relationships hold:

Lemma 3.3.7 *Both* $Z(t)$ *and* $\zeta(t)$ *are projections of* $\tilde{\zeta}(t)$ *onto their filtrations: for all* t,

- $Z(t) = \tilde{P}(\tilde{\zeta}(t) \mid \mathcal{F}_t)$,
- $\zeta(t) = \tilde{P}(\tilde{\zeta}(t) \mid \mathcal{G}_t)$.

Proof: We use the representation (3.4) of $\tilde{\zeta}(t)$:

$$\tilde{\zeta}(t) = \sum_{u\in N_t} \prod_{v<u} (1+A_v) e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \mathbf{1}_{(\xi_t=u)}. \qquad (3.13)$$

From this it follows that

$$\begin{aligned}\tilde{P}(\tilde{\zeta}(t)|\mathcal{F}_t) &= \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \times \prod_{v<u}(1+A_v)\,\tilde{P}(\mathbf{1}_{(\xi_t=u)}|\mathcal{F}_t) \\ &= \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) = Z(t),\end{aligned}$$

since $\tilde{P}(\mathbf{1}_{(\xi_t=u)}|\mathcal{F}_t) = \mathbf{1}_{(u\in N_t)} \times \prod_{v<u}(1+A_v)^{-1}$.

On the other hand, the martingale terms in (3.12) imply

$$\tilde{P}(\tilde{\zeta}(t)|\mathcal{G}_t) = \zeta(t) \times \tilde{P}\Big(\prod_{v<\xi_t}(1+A_v)e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}|\mathcal{G}_t\Big) = \zeta(t),$$

□

3.4 Changing the measures

In chapter 2 we used the single-particle martingale $\zeta_\lambda(t)$ to define new measures for the single-particle model, via

$$\left.\frac{\mathrm{d}\mathbb{P}_\lambda}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} = \frac{\zeta_\lambda(t)}{\zeta(0)}.$$

We have seen the close relationships between the three martingales ζ_λ, Z_λ and $\tilde{\zeta}_\lambda$:

$$Z_\lambda(t) = \tilde{P}\big(\tilde{\zeta}_\lambda(t)\,|\,\mathcal{F}_t\big), \qquad \zeta_\lambda(t) = \tilde{P}\big(\tilde{\zeta}_\lambda(t)\,|\,\mathcal{G}_t\big),$$

and in this section we show in a more general form how these close relationships mean that a new measure $\tilde{\mathbb{Q}}_\lambda$ defined in terms of $\tilde{P}$ as

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_t} = \frac{\tilde{\zeta}_\lambda(t)}{\tilde{\zeta}_\lambda(0)},$$

will induce measure changes on the sub-filtrations $\mathcal{G}_t$ and $\mathcal{F}_t$ of $\tilde{\mathcal{F}}_t$ whose Radon-Nikodym derivatives are given by $\zeta_\lambda(t)$ and $Z_\lambda(t)$ respectively.

We recall that in our set up we have a finest filtration $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ associated with the measure $\tilde{P}$, and two sub-filtrations $(\mathcal{F}_t)_{t\geq 0}$ with measure P and $(\mathcal{G}_t)_{t\geq 0}$ with measure $\mathbb{P}$. The martingale $\tilde{\zeta}$ can change the measure $\tilde{P}$:

Definition 3.4.1 *A measure $\tilde{\mathbb{Q}}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ is defined via its Radon-Nikodym derivative with respect to $\tilde{P}$:*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_t} = \frac{\tilde{\zeta}(t)}{\tilde{\zeta}(0)}.$$

Precisely, this notation means that for each event $F \in \tilde{\mathcal{F}}_t$ we define $\tilde{\mathbb{Q}}(F) := \tilde{P}\Big(\tilde{\zeta}(t)/\tilde{\zeta}(0); F\Big)$.

As we did for the measures P and $\mathbb{P}$, we can restrict $\tilde{\mathbb{Q}}$ to the sub-filtrations:

Definition 3.4.2 *We define the measure $\mathbb{Q}$ on $(\tilde{\mathcal{T}}, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0})$ via*

$$\mathbb{Q} := \tilde{\mathbb{Q}}|_{\mathcal{F}_\infty}.$$

Definition 3.4.3 *We define the measure $\hat{\mathbb{P}}$ on $(\tilde{\mathcal{T}}, \mathcal{G}_\infty, (\mathcal{G}_t)_{t\geq 0})$ via*

$$\hat{\mathbb{P}} := \tilde{\mathbb{Q}}|_{\mathcal{G}_\infty}.$$

A consequence of our new formulation in terms of filtrations and the equalities of Lemma 3.3.7 is that the changes of measure are carried out by $Z(t)$ and $\zeta(t)$ on their subfiltrations:

Theorem 3.4.4

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}P}\right|_{\mathcal{F}_t} = \frac{Z(t)}{Z(0)}, \qquad \textit{and} \qquad \left.\frac{\mathrm{d}\hat{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} = \frac{\zeta(t)}{\zeta(0)}.$$

Proof: These two results actually follow from a more general observation that if $\tilde{\mu}_1$ and $\tilde{\mu}_2$ are two measures defined on a measure space $(\Omega, \tilde{\mathcal{S}})$ with Radon-Nikodym derivative

$$\frac{\mathrm{d}\tilde{\mu}_2}{\mathrm{d}\tilde{\mu}_1} = f, \tag{3.14}$$

and if $\mathcal{S}$ is a sub-σ-algebra of $\tilde{\mathcal{S}}$, then the two measures $\mu_1 := \tilde{\mu}_1|_{\mathcal{S}}$ and $\mu_2 := \tilde{\mu}_2|_{\mathcal{S}}$ on $(\Omega, \mathcal{S})$ are related by the conditional expectation operation:

$$\frac{\mathrm{d}\mu_2}{\mathrm{d}\mu_1} = \tilde{\mu}_1(f|\mathcal{S}).$$

The proof of this is that if $g \in m\mathcal{S}$ and $S \in \mathcal{S}$ then

$$\begin{aligned}
\int_S g \,\mathrm{d}\mu_2 = \int_S g \,\mathrm{d}\tilde{\mu}_2 & \qquad \text{since } g \text{ is also in } m\tilde{\mathcal{S}}, \text{ and } S \in \mathcal{S} \text{ too},\\
= \int_S g\, f \,\mathrm{d}\tilde{\mu}_1 & \qquad \text{by (3.14)},\\
= \int_S \tilde{\mu}_1(gf|\mathcal{S}) \,\mathrm{d}\tilde{\mu}_1 & \qquad \text{by definition of the conditional expectation},\\
= \int_S g\, \tilde{\mu}_1(f|\mathcal{S}) \,\mathrm{d}\tilde{\mu}_1 & \qquad \text{since } g \text{ is } \mathcal{S}\text{-measurable},\\
= \int_S g\, \tilde{\mu}_1(f|\mathcal{S}) \,\mathrm{d}\mu_1 & \qquad \text{since everything is in } m\mathcal{S}.
\end{aligned}$$

Applying this general result (3.14) using the relationships between the general martingales given in Lemma 3.3.7 concludes the proof. □

3.4.1 Understanding the measure $\tilde{\mathbb{Q}}$

As the name suggests, we should be able to think of the spine as the backbone of the branching process. This is made precise by the following decomposition originally given by Chauvin and Rouault [8], and which now has become a standard part of most spine-based studies. We state this in a form taken from Kyprianou's spine-proofs for BBM [35]:

Theorem 3.4.5 *The measure $\tilde{P}$ on $\tilde{\mathcal{F}}_t$ can be decomposed as:*

$$\mathrm{d}\tilde{P}(\tau, M, \xi) = \mathrm{d}\mathbb{P}(\xi)\mathrm{d}\mathbb{L}^{(R(\xi_t))}(n_t) \prod_{v<\xi_t} \frac{1}{1+A_v} \prod_{v<\xi_t} p_{A_v}(\xi_{S_v}) \prod_{j=1}^{A_v} \mathrm{d}P\big((\tau, M)^v_j\big), \tag{3.15}$$

where $\mathbb{L}^{(R(\xi_t))}$ is the law of the Poisson (Cox) process with rate $R(\xi_t)$ at time t, and we remember that n_t counts the number of fission times on the spine before time t.

We can offer a reading of this decomposition, which we summarize as a lemma:

Lemma 3.4.6 *The terms of (3.15) respectively represent that under $\tilde{P}$,*

1. *the spine's motion is determined by the single-particle measure $\mathbb{P}$;*

2. *the fission times along the spine are independent of its motion, and occur as a Poisson process with a time-t rate of $R(\xi_t)$;*

3. *the spine is chosen uniformly from all the rays in the family tree;*

4. *at the fission time of node v on the spine, a random number A_v of children are born, with A_v being chosen independently and distributed according to the location-dependent random variable $A(\xi_{S_v})$;*

5. *each of these children gives rise to the subtrees $(\tau, M)_j^v$, for $1 \leq j \leq A_v$, which are not part of the spine and which are then determined by an independent copy of the original measure P shifted to their point and time of creation.*

This decomposition of $\tilde{P}_t$ given at (3.15) will allow us to interpret the measure $\tilde{\mathbb{Q}}$ if we appropriately factor the components of the change-of-measure martingale $\tilde{\zeta}(t)$ across this representation: on $\tilde{\mathcal{F}}_t$,

$$
\begin{aligned}
\mathrm{d}\tilde{\mathbb{Q}} &= \tilde{\zeta}(t)\,\mathrm{d}\tilde{P} && (3.16)\\
&= \zeta(t) \times e^{-\int_0^t R(\xi_s)\mathrm{d}s}\big(1 + m(\xi_{S_v})\big)^{n_t} \times \prod_{v<\xi_t} \frac{1+A_v}{1+m(\xi_{S_v})} \times \mathrm{d}\tilde{P} \\
&= \mathrm{d}\hat{\mathbb{P}}(\xi)\,\mathrm{d}\mathbb{L}^{((1+m(\xi_t))R(\xi_t))}(n) \prod_{v<\xi_t} \frac{1}{1+A_v} \prod_{v<\xi_t} \frac{1+A_v}{1+m(\xi_{S_v})} p_{A_v} \prod_{j=1}^{A_v} \mathrm{d}P\big((\tau, M)_j^v\big). && (3.17)
\end{aligned}
$$

Just as we did for $\tilde{P}$, we can offer a reading of this decomposition:

Lemma 3.4.7 *Under the measure $\tilde{\mathbb{Q}}$,*

1. *the spine process ξ_t moves as if under the* changed *measure $\hat{\mathbb{P}}$;*

2. *the birth times along the spine are independent of its motion and occur at an* accelerated *rate $(1 + m(\xi_t))R(\xi_t)$;*

3. *up until time t the nodes that make up the spine are still chosen uniformly;*

4. *at the birth time of node v on the spine, a random number A_v of subtrees $(\tau, M)_j^v$ (for $j = 1, \ldots, A_v$) are born, with A_v being chosen independently of the spine's motion and distributed according to the* tilted *distribution $((1+k)p_k(\xi_{S_v})/(1+m(\xi_{S_v})) : k = 0, 1, \ldots)$;*

5. *each of the subtrees which are not part of the spine still initiate P-multitype BBM from their space-type-time point of creation – and in this sense being unaffected by the change of measure.*

Such an interpretation of the measure $\tilde{\mathbb{Q}}$ was first given by Chauvin and Rouault [8] in the context of BBM, allowing them to come to the important conclusion that under the new measure $\tilde{\mathbb{Q}}$ the branching diffusion remains largely unaffected, except that the brownian particles of a single (random) line of descent in the family tree are given a changed motion, with an accelerated birth rate – they did not have random family sizes, so the size-biasing aspect was not seen. In the context of spines, size-biasing was first introduced in the Lyons *et al* papers [40, 34, 41].

3.4.2 A discussion of the martingales and new measures

In the preceding sections we have stated without proof that $Z(t)$ and $\tilde{\zeta}(t)$ are martingales, and we have used $\tilde{\zeta}(t)$ to define the measure $\tilde{\mathbb{Q}}$ on $(\tilde{\mathcal{T}},\tilde{\mathcal{F}}_\infty)$ which in turn has given us the measures $\mathbb{Q}$ on $(\tilde{\mathcal{T}},\mathcal{F}_\infty)$ and $\hat{\mathbb{P}}$ on $(\tilde{\mathcal{T}},\mathcal{G}_\infty)$.

In terms of rigour there are two issues with our approach: martingales must be proven to be martingales, and as we discussed in the earlier classical approach to the single-particle measure $\mathbb{P}_\lambda$ defined at Definition 2.3.5, martingales can be used as Radon-Nikodym derivatives to define measures on the σ-algebra $\tilde{\mathcal{F}}_t$ for all $t<\infty$, but we cannot use them to define a measure on $\tilde{\mathcal{F}}_\infty$ (since they are not in general *uniformly integrable*). We here explain how these both can be resolved easily.

First we consider the existence of the measure $\tilde{\mathbb{Q}}$ on $(\tilde{\mathcal{T}},\tilde{\mathcal{F}}_\infty)$. The existence of the original measure $\tilde{P}$ on $(\tilde{\mathcal{T}},\tilde{\mathcal{F}}_\infty)$, defined at Definition 3.2.1 is unproblematic since standard theory can be used (we referenced the work of Neveu [43] and Chauvin [7, 5]). In exactly the same way, Lemma 3.4.7 could actually be used to *define* $\tilde{\mathbb{Q}}$ on $(\tilde{\mathcal{T}},\tilde{\mathcal{F}}_\infty)$. From decomposition (3.17) and (3.16) it would then follow that $\tilde{\zeta}(t)$ is the Radon-Nikodym derivative between $\tilde{\mathbb{Q}}$ and $\tilde{P}$ on $\tilde{\mathcal{F}}_t$.

Thus $\tilde{\zeta}(t)$ is certainly a $(\tilde{\mathcal{T}},\tilde{\mathcal{F}}_t,\tilde{P})$-martingale, and it would then follow from Theorem 3.4.4 that $Z(t)$ is correctly a $(\tilde{\mathcal{T}},\mathcal{F}_t,P)$-martingale – as we initially stated in definition 3.10.

3.5 The spine decomposition

One of the most important results introduced in Lyons [40] was the so-called spine decomposition, which in the case of the additive martingale

$$Z_\lambda(t)=\sum_{u\in N_t} v_\lambda(Y_u(t))e^{\lambda X_u(t)-E_\lambda t},$$

from the finite-type branching diffusion of chapter 2 would be:

$$\tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty)=\sum_{u<\xi_t} v_\lambda(\eta_{S_u})e^{\lambda\xi_{S_u}-E_\lambda S_u}+v_\lambda(\eta_t)e^{\lambda\xi_t-E_\lambda t}. \tag{3.18}$$

To prove this we start by decomposing the martingale as

$$Z_\lambda(t)=\sum_{u\in N_t,u\notin\xi} v_\lambda(Y_u(t))e^{\lambda X_u(t)-E_\lambda t}+v_\lambda(\eta_t)e^{\lambda\xi_t-E_\lambda t},$$

which is clearly true since one of the particles $u\in N_t$ must be the in the line of descent that makes up the spine ξ. Recalling that the σ-algebra $\tilde{\mathcal{G}}_\infty$ contains all information about the line of nodes that makes up the spine, all about the spine diffusion (ξ_t,η_t) for all times t, and also contains all information regarding the fission times on the spine, it is useful to partition the particles $v\in\{u\in N_t,u\notin\xi\}$ into the distinct subtrees $(\tau,M)^u$ that were born at the fission times S_u from the particles that made up the spine before time t, or in other words those nodes in the $\{u<\xi_t\}$ of ancestors of the current spine node ξ_t. Thus:

$$Z_\lambda(t)=\sum_{u<\xi_t} e^{\lambda\xi_{S_u}-E_\lambda S_u}\left\{\sum_{v\in N_t,v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\right\}+v_\lambda(\eta_t)e^{\lambda\xi_t-E_\lambda t}.$$

If we now take the $\tilde{\mathbb{Q}}_\lambda$-conditional expectation of this, we find

$$\tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty) = v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t} + \sum_{u<\xi_t} e^{\lambda\xi_{S_u} - E_\lambda S_u}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{v\in N_t, v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\big|\,\tilde{\mathcal{G}}_\infty\Big).$$

We know from the decomposition (3.17) that the under the measure $\tilde{\mathbb{Q}}_\lambda$ the subtrees coming off the spine evolve as if under the measure P, and therefore

$$\begin{aligned}&\tilde{\mathbb{Q}}_\lambda\Big(\sum_{v\in N_t, v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\big|\,\tilde{\mathcal{G}}_\infty\Big)\\ &= \tilde{P}\Big(\sum_{v\in N_t, v\in(\tau,M)^u} v_\lambda(Y_v(t))e^{\lambda(X_u(t)-\xi_{S_u})-E_\lambda(t-S_u)}\big|\,\tilde{\mathcal{G}}_\infty\Big) = v_\lambda(\eta_{S_u}),\end{aligned}$$

since the additive expression being evaluated on the subtree is just a shifted copy of the martingale Z_λ itself.

This concludes the proof of (3.18), but before we go move on to give a similar proof for the general case, for easier reference through the cumbersome-looking general proof it is worth recalling that

$$\zeta_\lambda(t) = e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t},$$

and therefore noting that (3.18) can alternatively be written as

$$\tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty) = \sum_{u<\xi_t} e^{-\int_0^{S_u} R(\eta_s)\,\mathrm{d}s}\zeta_\lambda(S_u) + e^{-\int_0^t R(\eta_s)\,\mathrm{d}s}\zeta_\lambda(t).$$

Also, in the general model we are supposing that each particle u in the spine will give birth to a total of A_u subtrees that go off from the spine – the one remaining other offspring is used to continue the line of descent that makes up the spine. This explains the appearance of A_u in the decomposition.

Theorem 3.5.1 (Spine decomposition) *We have the following* ***spine decomposition*** *for the additive branching-particle martingale:*

$$\tilde{\mathbb{Q}}^x(Z(t)|\tilde{\mathcal{G}}_\infty) = \sum_{u<\xi_t} A_u\, e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(S_u) + e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t).$$

Proof: In each sample tree one and only one of the particles alive at time t is the spine and therefore:

$$\begin{aligned}Z(t) &= \sum_{u\in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t),\\ &= e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t) + \sum_{u\in N_t, u\neq\xi_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t).\end{aligned}$$

The other individuals $\{u \in N_t, u \neq \xi_t\}$ can be partitioned into subtrees created from fissions along the spine. That is, each node u in the spine ξ_t (so $u < \xi_t$) has given birth at time S_u to one offspring node uj (for some $1 \leq j \leq 1 + A_u$) that was chosen to continue the spine whilst the other A_u individuals go off to make the subtrees $(\tau, M)^u_j$. Therefore,

$$Z(t) = e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t) + \sum_{u<\xi_t} e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s} \sum_{\substack{j=1,\dots,1+A_u \\ uj\notin\xi}} Z_{uj}(S_u;t), \tag{3.19}$$

where for $t \geq S_u$,

$$Z_{uj}(S_u;t) := \sum_{v\in N_t, v\in(\tau,M)^u_j} e^{-\int_{S_u}^t m(X_v(s))R(X_v(s))\,\mathrm{d}s}\zeta_v(t),$$

is, conditional on $\tilde{\mathcal{G}}_\infty$, a $\tilde{P}$-martingale on the subtree $(\tau, M)^u_j$, and therefore

$$\tilde{P}\big(Z_{uj}(S_u;t)|\tilde{\mathcal{G}}_\infty\big) = \zeta(S_u).$$

Thus taking $\tilde{\mathbb{Q}}$-conditional expectations of (3.19) gives

$$\begin{aligned}
\tilde{\mathbb{Q}}^x\big(Z(t)|\tilde{\mathcal{G}}_\infty\big) &= e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t) + \sum_{u<\xi_t} e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s}\tilde{P}\Big(\sum_{\substack{j=1,\dots,1+A_u \\ uj\notin\xi}} Z_{uj}(S_u;t), \big|\, \tilde{\mathcal{G}}_\infty\Big), \\
&= e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(t) + \sum_{u<\xi_t} e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s}A_u\,\zeta(S_u),
\end{aligned}$$

which completes the proof. □

This representation was first used in the Lyons *et al* [40, 34, 41] papers and has become the standard way to investigate the behaviour of Z under the measure $\tilde{\mathbb{Q}}$. We observe that the two measures $\tilde{P}$ and $\tilde{\mathbb{Q}}$ for the general model are equal when conditioned on $\tilde{\mathcal{G}}_\infty$ since this factors out their differences in the spine diffusion ξ_t, the family sizes born from the spine and the fission times on the spine. Therefore it follows that the same argument as used above applies for $\tilde{P}$ to give:

Corollary 3.5.2

$$\tilde{P}\big(Z(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{u<\xi_t} A_u\, e^{-\int_0^{S_u} m(\xi_s)R(\xi_s)\,\mathrm{d}s}\zeta(S_u) + e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}\tilde{P}^{\xi_{S_u}}\big(\zeta(0)\big).$$

As a matter of fact, this above representation for $\tilde{P}\big(Z(t)|\tilde{\mathcal{G}}_\infty\big)$ has not received any attention in the literature, and in the context of our particular finite-type branching diffusion we can use it to explain a subtle difference between the spine and branching pictures. We note that the following remarks are not needed elsewhere in the thesis and are just to be seen as a side observation.

We proved at Lemma 3.3.7 that the martingales Z_λ and ζ_λ are both projections of the martingale $\tilde{\zeta}_\lambda$:

$$\begin{aligned}
Z_\lambda(t) &= \tilde{P}\big(\tilde{\zeta}_\lambda(t)\,|\,\mathcal{F}_t\big), \quad \text{and} \\
\zeta_\lambda(t) &= \tilde{P}\big(\tilde{\zeta}_\lambda(t)\,|\,\mathcal{G}_t\big).
\end{aligned}$$

Given that $Z(t) \in m\mathcal{F}_t \subset m\tilde{\mathcal{F}}_t$, it is natural to wonder if we could also have the equality $\zeta_\lambda(t) = \tilde{P}\big(Z_\lambda(t)\,|\,\mathcal{G}_t\big)$?

Since the algebra $\mathcal{G}_t$ is not a sub-algebra of $\mathcal{F}_t$, this is not immediately a trivial question, and in fact the result *does not hold* – basically because the branching process conditioned on the spine contains more information than the spine itself. To see this we first note that $\mathcal{G}_t$ is a sub-algebra of $\tilde{\mathcal{G}}_\infty$, whence by the tower property

$$\tilde{P}\big(Z_\lambda(t)\,|\,\mathcal{G}_t\big) = \tilde{P}\big(\tilde{P}(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty)\,|\,\mathcal{G}_t\big).$$

At the same time, since $\zeta_\lambda(t) = \tilde{P}\big(\tilde{\zeta}_\lambda|\mathcal{G}_t\big)$, it follows that the equality $\zeta_\lambda(t) = \tilde{P}\big(Z_\lambda(t)\,|\,\mathcal{G}_t\big)$ is equivalent to showing

$$\tilde{P}\big(\tilde{\zeta}_\lambda|\mathcal{G}_t\big) = \tilde{P}\big(\tilde{P}(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty)\,|\,\mathcal{G}_t\big),$$

which by definition means showing that for all $G \in \mathcal{G}_t$ we have the equality

$$\int_G \tilde{\zeta}_\lambda(t)\,\mathrm{d}\tilde{P} = \int_G \tilde{P}\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big)\,\mathrm{d}\tilde{P}. \tag{3.20}$$

However, just by comparing

$$\tilde{\zeta}_\lambda(t) = \prod_{u<\xi_t}(1+A_v) \times v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t},$$

with

$$\tilde{P}\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{u<\xi_t} v_\lambda(\eta_{S_u})e^{\lambda\xi_{S_u} - E_\lambda S_u} + v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t}$$

it is clear that $\tilde{P}\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big)$ makes references to the spine's position at the earlier times S_u, whilst $\tilde{\zeta}_\lambda(t)$ does not. Thus if the event G depends on the history of the spine up to time t then likely as not $\tilde{\zeta}_\lambda(t)$ and $\tilde{P}\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big)$ will be different, and equality will not hold.

What is happening here is that the births of subtrees on the spine act as a memory of where the spine has been in the past, so that conditional on knowing which particle is currently the spine, the algebra $\mathcal{F}_t$ contains *more information* about the spine than $\mathcal{G}_t$, and therefore

$$\zeta_\lambda(t) \neq \tilde{P}\big(Z_\lambda(t)\,|\,\mathcal{G}_t\big).$$

3.6 New spine results

Having covered the formal basis for our spine approach, we now present some new results that follow from our spine formulation: the Gibbs-Boltzmann weights, conditional expectations, and a simpler proof of the improved Many-to-One theorem.

3.6.1 The *Gibbs-Boltzmann* weights of $\tilde{\mathbb{Q}}$

The Gibbs-Boltzmann weightings in branching processes are well-known, for example see Chauvin and Rouault [6] where they consider random measures on the boundary of the tree, and Harris [24] which gives convergence results for Gibbs-Boltzmann random measures. They have previously been considered via the individual terms of the additive martingale Z, but the following theorem gives a new interpretation of these weightings in terms of the spine. We recall that

$$Z(t) = \sum_{u \in N_t} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t).$$

Theorem 3.6.1 *Let $u \in \Omega$ be a given and fixed label. Then*

$$\tilde{\mathbb{Q}}(\xi_t = u | \mathcal{F}_t) = \mathbf{1}_{(u \in N_t)} \frac{e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t)}{Z(t)}.$$

Proof: Suppose $u \in \Omega$, and $F \in \mathcal{F}_t$. We aim to show:

$$\int_F \mathbf{1}_{(\xi_t = u)} \,\mathrm{d}\tilde{\mathbb{Q}}(\tau, M, \xi) = \int_F \mathbf{1}_{(u \in N_t)} \frac{e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t)}{Z(t)} \,\mathrm{d}\tilde{\mathbb{Q}}(\tau, M, \xi).$$

First of all we know that $\mathrm{d}\tilde{\mathbb{Q}}/\mathrm{d}\tilde{P} = \tilde{\zeta}(t)$ on $\mathcal{F}_t$ and therefore,

$$\text{LHS} = \int_F \mathbf{1}_{(\xi_t = u)} \prod_{v < \xi_t} (1 + A_v) e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \zeta(t) \,\mathrm{d}\tilde{P}(\tau, M, \xi),$$

by definition of $\tilde{\zeta}(t)$ at (3.12). The definition 3.2.3 of the measure $\tilde{P}$ requires us to express the integrand with a representation like (3.4):

$$\begin{aligned}
&\mathbf{1}_{(\xi_t = u)} \prod_{v < \xi_t} (1 + A_v) e^{-\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s} \zeta(t) \\
&= \mathbf{1}_{(\xi_t = u)} \sum_{w \in N_t} \prod_{v < w} (1 + A_v) e^{-\int_0^t m(X_w(s))R(X_w(s))\,\mathrm{d}s} \zeta_w(t) \mathbf{1}_{(\xi_t = w)}, \\
&= \mathbf{1}_{(\xi_t = u)} \mathbf{1}_{(u \in N_t)} \prod_{v < u} (1 + A_v) e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t),
\end{aligned}$$

and therefore

$$\begin{aligned}
\text{LHS} &= \int_F \mathbf{1}_{(u \in N_t)} \prod_{v < u} (1 + A_v) e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \mathbf{1}_{(\xi_t = u)} \,\mathrm{d}\tilde{P}(\tau, M, \xi), \\
&= \int_F \mathbf{1}_{(u \in N_t)} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \,\mathrm{d}P(\tau, M, \xi),
\end{aligned}$$

by definition 3.2.3. We emphasize that now this is an integral taken with respect to the measure P over the σ-algebra $\mathcal{F}_t$, and here we know that $\mathrm{d}P/\mathrm{d}\mathbb{Q} = 1/Z(t)$, so:

$$\text{LHS} = \int_F \mathbf{1}_{(u \in N_t)} e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) \,\frac{1}{Z(t)} \mathrm{d}\mathbb{Q}(\tau, M, \xi),$$

and the proof is concluded. □

Comment

At first it may seem that result 3.6.1 could give rise to a paradox: suppose that two particles are at the same spatial position: $X_{u_1}(t) = X_{u_2}(t)$ but that u_1 has had more ancestors. According to the above there is an equal conditional probability that either is the spine, but we know that under $\tilde{\mathbb{Q}}$ the spine has an *increased* birth rate, and might therefore expect u_1 to have a higher probability of being the spine.

The way to resolve this is to realise that the spine is always chosen according to the harmonic measure on the tree and therefore

$$\tilde{\mathbb{Q}}(\xi_t = u_1 | N_t) = \mathbf{1}_{(u_1 \in N_t)} \prod_{v < u_1} (1 + A_v)^{-1},$$

which in fact gives u_1 a *smaller* probability of being the spine. This neutralizes the accelerated birth rate of the spine, and the two effects balance each other exactly.

3.6.2 Conditional expectations and the Kesten-Stigum Theorem

The above result combines with the representation (3.4) to show how we take conditional expectations under the measure $\tilde{\mathbb{Q}}$.

Theorem 3.6.2 *If $f(t) \in m\tilde{\mathcal{F}}_t$, and $f = \sum_{u \in N_t} f_u(t) \mathbf{1}_{(\xi_t = u)}$, with $f_u(t) \in m\mathcal{F}_t$ then*

$$\tilde{\mathbb{Q}}(f(t)|\mathcal{F}_t) = \sum_{u \in N_t} f_u(t) \frac{e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t)}{Z(t)}. \tag{3.21}$$

Proof: It is clear that

$$\tilde{\mathbb{Q}}(f(t)|\mathcal{F}_t) = \sum_{u \in N_t} f_u(t) \tilde{\mathbb{Q}}(\xi_t = u | \mathcal{F}_t),$$

and the result follows from Theorem 3.6.1. □

A simple corollary of this is exceptionally useful, and goes an awful long way to obtaining the Kesten-Stigum result in very general models (see [29, 28, 30] for the classical proofs of the Kesten-Stigum results),

Corollary 3.6.3 *If $g(\cdot)$ is a measurable function on J then*

$$\sum_{u \in N_t} g(X_u(t))\, e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s} \zeta_u(t) = \tilde{\mathbb{Q}}(g(\xi_t)|\mathcal{F}_t) \times Z(t).$$

Proof: It is easy to show that $g(\xi_t) = \sum_{u \in N_t} g(X_u(t)) \mathbf{1}(\xi_t = u)$, and now the result follows from the above theorem. □

In the case of BBM this particular corollary would state that for any Lebesgue-measurable function f we have

$$\sum_{u \in N_t} f(X_u(t)) e^{\lambda X_u(t) - (\frac{1}{2}\lambda^2 + r)t} = \tilde{\mathbb{Q}}_\lambda(f(\xi_t)|\mathcal{F}_t) \times Z_\lambda(t),$$

The Kesten-Stigum theorem gives conditions under which an operation like this converge as $t \to \infty$, and in general it is found that when it exists the limit is a multiple of the martingale limit $Z_\lambda(\infty)$ – see Lyons *et al* [34] for a proof of this based on other spine techniques. Our improved spine formulation therefore gives a previously unknown but simple meaning to this operation, in terms of a conditional expectation. Furthermore, this new interpretation might also give good grounds for alternative spine proofs of the Kesten-Stigum theorem, via a study of when the conditional expectation $\tilde{\mathbb{Q}}(g(\xi_t)|\mathcal{F}_t)$ converges.

Remark: The convergence of conditional expectations is not necessarily easy to obtain and has not yet been considered to any great extent in the literature. The *Gibbs Conditioning Principle* (see Dembo and Zeitouni [12]) is one useful example of a result for conditional expectations, and it may be possible to adapt it to this context but this is left to further research.

3.6.3 The *Full* Many-to-One Theorem

Much of the work of Harris and Williams [21] and Harris and Git [25] made important use of a *Many-to-One* result, which had been proved using resolvents (see Champneys *et al* [4] for a similar proof). We recall our statement in chapter 2 of the Many-to-One theorem for the finite-type model:

Theorem 3.6.4 *For any measurable function $f : J \to \mathbb{R}$ we have*

$$P^{x,y}\Big(\sum_{u\in N_t} f(X_u(t), Y_u(t))\Big) = \mathbb{P}^{x,y}\Big(e^{\int_0^t R(\eta_s)\,\mathrm{d}s} f(\xi_t, \eta_t)\Big).$$

Intuitively it is clear that the up-weighting term $e^{\int_0^t R(\eta_s)\,\mathrm{d}s}$ incorporates the notion of the population growing at an exponential rate, whilst the idea of $f(\xi_t, \eta_t)$ being the 'typical' behaviour of $f(X_u(t), Y_u(t))$ is also reasonable.

The main problem with the proof given by Harris and Williams is that it applies only to functions of the above form that therefore depend only on *the time-t location* of the spine – it does not cover functions that depend on the entire *path history* of the spine up to time t.

With the spine approach we have the benefit of being able to give a much less complicated proof of the stronger version that covers the most general path-dependent functions.

Theorem 3.6.5 (Many-to-One) *If $g(t) \in m\mathcal{G}_t$ has the representation*

$$g(t) = \sum_{u\in N_t} g_u(t)\mathbf{1}_{(\xi_t = u)},$$

where $g_u(t) \in m\mathcal{F}_t$, then

$$P\Big(\sum_{u\in N_t} g_u(t)\Big) = \mathbb{P}\Big(e^{\int_0^t m(\xi_s)R(\xi_s)\mathrm{d}s} g(t)\Big).$$

Proof: Let $f(t)$ be any given $\mathcal{G}_t$-measurable function. Since $\mathcal{G}_t \subset \tilde{\mathcal{F}}_t$ it therefore follows that $f(t)$ is also $\tilde{\mathcal{F}}_t$-measurable and we can use the tower property together with Theorem 3.6.2 to obtain

$$\begin{aligned}\tilde{\mathbb{Q}}(f(t)) &= \tilde{\mathbb{Q}}\Big(\tilde{\mathbb{Q}}(f(t)|\mathcal{F}_t)\Big) = \mathbb{Q}\Big(\tilde{\mathbb{Q}}(f(t)|\mathcal{F}_t)\Big) \\ &= \mathbb{Q}\Big(\frac{1}{Z(t)}\sum_{u\in N_t} f_u(t)e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)\Big).\end{aligned}$$

We emphasize that this is a $\mathbb{Q}$ expectation of a $\mathcal{F}_t$-measurable expression. Theorem 3.4.4 states that

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}P}\right|_{\mathcal{F}_t} = \frac{Z(t)}{Z(0)},$$

and therefore we conclude from the above that

$$\tilde{\mathbb{Q}}(f(t)) = Z(0)^{-1}P\Big(\sum_{u\in N_t} f_u(t)e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)\Big).$$

On the other hand, since $f(t)$ is $\mathcal{G}_t$-measurable, Theorem 3.4.4 states that

$$\left.\frac{\mathrm{d}\hat{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_t} = \frac{\zeta(t)}{\zeta(0)},$$

and from the definitions 3.2.6 and 3.4.3 this means

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}}{\mathrm{d}\tilde{P}}\right|_{\mathcal{G}_t} = \frac{\zeta(t)}{\zeta(0)},$$

and therefore

$$\tilde{\mathbb{Q}}(f(t)) = \zeta(0)^{-1}\tilde{P}(f(t)\times\zeta(t)).$$

Trivially we always have $Z(0) = \zeta(0)$ and thus we have shown in the first instance that for any $f(t) \in m\mathcal{G}_t$,

$$P\Big(\sum_{u\in N_t} f_u(t)e^{-\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)\Big) = \tilde{P}(f(t)\times\zeta(t)). \tag{3.22}$$

Given $g(t) \in m\mathcal{G}_t$, we can define

$$f(t) := e^{\int_0^t m(\xi_s)R(\xi_s)\,\mathrm{d}s}g(t)\times\zeta(t)^{-1},$$

which is clearly $\mathcal{G}_t$-measurable and satisfies $f(t) = \sum_{u\in N_t} f_u(t)\zeta_u(t)$ with

$$f_u(t) = g_u(t)e^{\int_0^t m(X_u(s))R(X_u(s))\,\mathrm{d}s}\zeta_u(t)^{-1} \in m\mathcal{F}_t.$$

When we now use this $f(t)$ in the above (3.22) we arrive at the result to be proved. □

In the case in which $g = g(\xi_t)$ for some Borel-measurable function $g(\cdot)$, the trivial representation

$$g(\xi_t) = \sum_{u\in N_t} g(X_u(t))\mathbf{1}_{(\xi_t=u)}$$

leads immediately to the weaker version of the Many-to-One result that was originally proven by Harris and Williams using resolvents and the Feynman-Kac formula, expressed in terms of our more general branching Markov process $\mathbb{X}_t$:

Corollary 3.6.6 *If $g(\cdot): J\to\mathbb{R}$ is $\mathcal{B}$-measurable then*

$$P\Big(\sum_{u\in N_t} g(X_u(t))\Big) = \mathbb{P}\Big(e^{\int_0^t R(\xi_s)\mathrm{d}s}g(\xi_t)\Big).$$

Chapter 4

Spine proofs for $\mathcal{L}^p$-convergence

4.1 Overview

We are going to use spine techniques to consider the $\mathcal{L}^p$-convergence properties (for $p > 1$) of the strictly-positive martingales for three different models of branching diffusions. It is a common feature of these diffusion models, where there is actually a family of such martingales $\{Z_\lambda : \lambda \in \mathbb{R}\}$, that for all λ within an open interval about 0 the martingale Z_λ is convergent in $\mathcal{L}^p$ for some $p > 1$; for λ outside of this interval the limit of Z_λ is almost surely null. For values of λ at the boundary of this interval, the so-called *critical* values, it has been conjectured (and in some models proven) that the martingales have a null limit – we give a proof for the simpler first model but not for the others since they require different techniques using 'derivative' martingales – see Kyprianou [35] or Harris [23] for examples.

The first model is branching Brownian motion (BBM), for which a proof of $\mathcal{L}^p$-convergence was originally given by Neveu [44] using classical techniques based on the branching decomposition. We actually deal with two variants of BBM: first sections deal with the case of binary-splitting where the fissions produce only two particles, and in section 4.2.3 we extend the model to allow the fissions to produce a *random* number of offspring.

The second model we look at is the finite-type branching diffusion of chapter 2 and is a generalization of a 2-type model treated in the paper by Champneys *et al* [4]. Again, we deal separately with the cases of binary-splitting and random family sizes, and in this second case allow the distribution of the family sizes to be type-dependent.

The third model has a *continuous* type-space where the type of each particle moves independently as an Orstein-Uhlenbeck process on $\mathbb{R}$. This branching diffusion was first introduced in Harris and Williams [21] and has also been investigated in Harris [22], Harris and Git [25] and Kyprianou and Engländer [14].

We remark that not all the results we prove in this chapter are new, since some have already been proven using classical techniques: references are Neveu [44] for BBM (but not for the random-family sizes model), and Harris and Williams [21] for the continuous-type model. And of course we have already seen our classical approach for the finite-type model in chapter

2 (but the result for type-dependent random family sizes is new). In all cases spine proofs offer a much simpler, more intuitive and consistent approach than the classical counterparts, especially for the more complex model of Harris and Williams [21]. Spine proofs for each of these models each run along similar lines, and it is a credit to the spine approach that this is possible. More classical techniques based on the expectation semigroup are simply not able to generalize easily, since they often require either some *a priori* bounds on the semigroup or involve difficult estimates – for example, in Harris and Williams [21] their important bound of a non-linear term is made possible only by the existence of a good $\mathcal{L}^2$ theory for their operator, and this is not generally available.

As mentioned, there are a number of reasons why we may be interested in knowing about the $\mathcal{L}^p$ convergence of a martingale, but for this thesis we would particularly draw attention to the *techniques* that we use since similar ideas are the basis of proofs of two important problems in the large-deviations theory of branching diffusions in the second part of this thesis.

Briefly, to prove that the martingale converges in $\mathcal{L}^p$ for some $p > 1$ we use Doob's theorem, and therefore need only to show that the martingale is *bounded* in $\mathcal{L}^p$. The *spine decomposition* is an excellent tool here for showing boundedness of the martingale since it reduces difficult calculations over the whole collection of branching particles to just the single spine process. When $\mathcal{L}^p$-convergence does not hold the martingale limit is almost-surely null, and we prove this by showing that the martingale is almost-surely *not bounded* in a new measure – this approach relies on a measure-theoretic result given below and has become standard in the spine methodology since the important work of Lyons *et al* [41, 40, 34]. We show unboundedness of the martingale just by considering the contribution of the spine, which is shown to be unbounded.

4.2 Branching Brownian motion

Consider a branching Brownian motion (BBM) with constant branching rate r, which is the branching process whereby particles diffuse independently according to a Brownian motion and at any moment undergo fission at a rate r to produce two particles. We suppose that the probabilities of this are $\{P^x : x \in \mathbb{R}\}$ so that P^x is a measure defined on the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ such that it is the law of the process initiated from a single particle positioned at x.

Suppose that the configuration of this branching Brownian motion at time t is given by the $\mathbb{R}$-valued point process $\mathbb{X}_t := \{X_u(t) : u \in N_t\}$ where N_t is the set of individuals alive at time t. It is well known that for any $\lambda \in \mathbb{R}$,

$$Z_\lambda(t) := \sum_{u \in N_t} e^{-rt} e^{\lambda X_u(t) - \frac{1}{2}\lambda^2 t} \tag{4.1}$$

defines a *strictly-positive* P-martingale, so $Z_\lambda(\infty) := \lim_{t\to\infty} Z_\lambda(t)$ is almost surely finite under P^x.

4.2.1 The $\mathcal{L}^1$-convergence theorem of Kyprianou

We are going to use a change of measure together with the spine decomposition to determine the conditions under which this martingale (4.1) is $\mathcal{L}^p$-convergent for some $p > 1$, but first review a recent spine-proof given by Kyprianou [35] on the question of $\mathcal{L}^1$-convergence.

In fact Kyprianou [35] deals with a slightly more general version of the BBM model in which particle fissions may produce a *random* number of offspring (but always at least one), with that number chosen independently of a particle's position and distributed according to some given distribution on the integers; we look at this more general model in section 4.2.3. If we were only to consider the binary-splitting case, his result would be stated as follows:

Theorem 4.2.1 (Kyprianou) *Let* $\tilde{\lambda} := \sqrt{2r}$.

- *If* $|\lambda| < \tilde{\lambda}$ *then* $Z_\lambda(\infty) > 0$ *almost-surely and* $Z_\lambda(t) \to Z_\lambda(\infty)$ *in* $\mathcal{L}^1(P)$;
- *If* $|\lambda| \geq \tilde{\lambda}$ *then* $Z_\lambda(\infty) = 0$ *almost surely.*

The proofs for both parts of this theorem depend on a change-of-measure argument, and we briefly cover this now before reviewing Kyprianou's spine-proofs of the above theorem. We shall use similar measure-change and spine ideas in our later work on $\mathcal{L}^p$-convergence of the martingales.

The following measure-theoretic result explains how it is possible to deduce the convergence properties of a strictly-positive martingale when it is used to change measures as a Radon-Nikodym derivative, and sets the scene for the spine approach.

Theorem 4.2.2 *Suppose that* P *and* $\mathbb{Q}$ *are two probability measures on an underlying filtered space* $(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0})$, *such that for some strictly-positive martingale* Z_t,

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}P}\right|_{\mathcal{F}_t} = Z_t.$$

If we define $Z_\infty := \limsup_{t\to\infty} Z_t$, *then for any* $F \in \mathcal{F}_\infty$

$$\mathbb{Q}(F) = \int_F Z_\infty \,\mathrm{d}P + \mathbb{Q}(F \cap \{Z_\infty = \infty\}), \tag{4.2}$$

and consequently,

1. $\mathbb{Q}$ *is absolutely continuous with respect to* P *if and only if* $\int_\Omega Z_\infty \,\mathrm{d}P = 1$ *if and only if* $\mathbb{Q}(Z_\infty = \infty) = 0$;
2. $\mathbb{Q}$ *is singular with respect to* P *if and only if* $P(Z_\infty = 0) = 1$ *if and only if* $\mathbb{Q}(Z_\infty = \infty) = 1$.

A proof of the decomposition (4.2) can be found in Durrett [13], at page 241. Thus the aim is to determine a measure $\mathbb{Q}^x_\lambda$ for the BBM process $\mathbb{X}_t$ such that

$$\left.\frac{\mathrm{d}\mathbb{Q}^x_\lambda}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)}. \tag{4.3}$$

We note that the denominator $Z_\lambda(0)$ is a normalizing factor, and we make it explicit that (4.3) means that for all $F \in \mathcal{F}_t$ we have $\mathbb{Q}_\lambda^x(F) = P^x(Z_\lambda(t)/Z_\lambda(0); F)$.

It was work originally carried out by Chauvin and Rouault [8] that gave a pathwise construction of the measure $\mathbb{Q}_\lambda^x$ that would result in the relationship (4.3). The measure used by Kyprianou also involves another feature of size-biasing the family distributions, but we shall leave that until the section in which we look more carefully at models with random family sizes.

Definition 4.2.3 *We suppose that $\mathbb{Q}_\lambda^x$ is a measure such that under $\mathbb{Q}_\lambda^x$ the point process $\mathbb{X}_t$ evolves as follows:*

- *starting from position x, the original ancestor diffuses according to a Brownian motion on $\mathbb{R}$ with drift λ;*
- *at rate $2r$ the particle undergoes fission producing two particles;*
- *with equal probability, one of these two particles is selected;*
- *this chosen particle repeats stochastically the behaviour of the parent;*
- *the other particle initiates, from its birth position, an independent copy of a $P^{\cdot}$ branching Brownian motion with branching rate r.*

In this construction, the individuals that are selected to have a drift of λ make up a (random) line of descent which has come to be referred to as the *spine*, and we refer to the diffusion path that they generate as $(\xi_t)_{t\geq 0}$; below we refer to the times at which the spine undergoes fission as $(S_1, S_2, \ldots, S_{n_t})$, so that n_t is the number of fissions on the spine before time t and is therefore a Poisson process of rate $2r$ under $\mathbb{Q}_\lambda$. As shown by Chauvin and Rouault [8], we have the following important result:

Theorem 4.2.4 *The change of measure is given by:*

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^x}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = e^{-\lambda x} Z_\lambda(t). \tag{4.4}$$

This pathwise construction of the measure $\mathbb{Q}_\lambda$ is equivalent to the measures we constructed in the previous chapter, and we mention that we can assume that an extended measure $\tilde{\mathbb{Q}}_\lambda^x$ gives $\mathbb{X}_t$ the above behaviour and also can measure events that depend on the spine.

Proof of Part 1 of Theorem 4.2.1 Suppose that the σ-algebra $\tilde{\mathcal{G}}_\infty$ is generated by the diffusion path $(\xi_t)_{t\geq 0}$ of all the particles that make up the spine and the fission times $(S_1, S_2, \ldots)$. Under the measure $\tilde{\mathbb{Q}}_\lambda$ the particles that are not selected for the spine behave as if under the original measure P, and therefore it follows that

$$\tilde{\mathbb{Q}}(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty) = \sum_{k=1}^{n_t} e^{\lambda \xi_{S_k} - E_\lambda S_k} + e^{\lambda \xi_t - E_\lambda t}, \tag{4.5}$$

since the contributions from the sub-trees that grow out of the spine at the kth fission-time S_k on the spine are P-martingales and therefore are equal to their immediate contribution on being born. This is the spine decomposition that we discussed in the previous chapter.

Kyprianou [35] is now able to show that when $|\lambda| < \tilde{\lambda}$ the drift on the spine diffusion is such that

$$\limsup_{t\to\infty} \tilde{\mathbb{Q}}(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty) < \infty \qquad \tilde{\mathbb{Q}}\text{-a.s.},$$

and therefore with an application of Fatou's lemma,

$$\begin{aligned}\tilde{\mathbb{Q}}(\liminf_{t\to\infty} Z_\lambda(t)|\tilde{\mathcal{G}}_\infty) &\le \liminf_{t\to\infty} \tilde{\mathbb{Q}}(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty)\\ &\le \limsup_{t\to\infty} \tilde{\mathbb{Q}}(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty)\\ &< \infty.\end{aligned}$$

This must imply that

$$\liminf_{t\to\infty} Z_\lambda(t) < \infty \qquad \tilde{\mathbb{Q}}\text{-a.s.} \tag{4.6}$$

which is to say

$$\liminf_{t\to\infty} Z_\lambda(t) < \infty \qquad \mathbb{Q}\text{-a.s.}$$

since $\limsup_{t\to\infty} Z_\lambda(t)$ is $\mathbb{Q}$-measurable. Finally, we know that $1/Z_\lambda(t)$ is a strictly-positive $\mathbb{Q}$-martingale whence it is convergent almost-surely under $\mathbb{Q}$ and therefore

$$\infty > \liminf_{t\to\infty} Z_\lambda(t) = \lim_{t\to\infty} Z_\lambda(t) \qquad \mathbb{Q}\text{-a.s.}$$

Now the first part of Theorem 4.2.2 clinches it, giving

$$\mathbb{E}_P Z_\lambda(\infty) = 1.$$

In a Corollary 2.5.3 we showed how this implies that the martingale Z_λ is actually uniformly integrable, and therefore $\mathcal{L}^1$-convergent. □

Proof of Part 2 of Theorem 4.2.1 In order to show that the martingale limit of our $Z_\lambda(t)$ is null we use the second part of Theorem 4.2.2 and therefore intend to show that

$$\mathbb{Q}_\lambda\Big(\limsup_{t\to\infty} Z_\lambda(t) = \infty\Big) = 1.$$

Because one of the individuals $u \in N_t$ *must* be the spine, it is immediate that

$$Z_\lambda(t) = \sum_{u\in N_t} e^{\lambda X_u(t) - E_\lambda t} > e^{\lambda\xi_t - E_\lambda t} = e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t - rt}.$$

Under the measure $\tilde{\mathbb{Q}}_\lambda$ the spine has a linear drift equal to λ, whence we can write

$$e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t - rt} = e^{\lambda B_t + (\frac{1}{2}\lambda^2 - r)t},$$

where B_t is a $\tilde{\mathbb{Q}}_\lambda$-Brownian motion. Thus $\lambda^2 \ge 2r$ will force $\limsup_{t\to\infty} Z_\lambda(t) = \infty$ almost surely under $\tilde{\mathbb{Q}}_\lambda$. With Theorem 4.2.2 this completes the proof.

□

4.2.2 A spine proof of $\mathcal{L}^p$-convergence

We started out supposing that the probabilities of the BBM are $\{P^x : x \in \mathbb{R}\}$, so that P^x is a measure defined on the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ such that it is the law of the process initiated from a single particle positioned at x. Using the process we explained in the definition 3.2.3, we extend these measures to $\tilde{P}^x$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ by defining

$$\int_{\tilde{\mathcal{T}}} f \, \mathrm{d}\tilde{P}^x := \int_{\mathcal{T}} \sum_{u\in N_t} f_u \Big(\prod_{v<u} \frac{1}{2} \Big) \, \mathrm{d}P^x,$$

for each $f \in m\tilde{\mathcal{F}}_t$ with representation like (3.4). This $\tilde{P}^x$ is an extension to the original measure: $P^x = \tilde{P}^x|_{\mathcal{F}_\infty}$, and we have seen the general proof for this at Theorem 3.2.4.

Rather than using the pathwise construction of definition 4.2.3 we have seen in that in our formulation of the spine approach we can directly define $\tilde{\mathbb{Q}}_\lambda$ via a Radon-Nikodym derivative with respect to $\tilde{P}$:

Definition 4.2.5 *$\tilde{\mathbb{Q}}_\lambda$ is a measure on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ defined as:*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}^x_\lambda}{\mathrm{d}\tilde{P}^x}\right|_{\tilde{\mathcal{F}}_t} = e^{-\lambda x}\, 2^{n_t} e^{-rt} e^{\lambda \xi_t - \frac{1}{2}\lambda^2 t}. \tag{4.7}$$

From this definition it is clear that the martingale term $e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}$ is going to give the spine process ξ_t a drift of λ under $\tilde{\mathbb{Q}}_\lambda$. The other term $e^{-rt}2^{n_t}$ is actually a martingale term that will increase the rate of the Poisson process n_t of fission times on the spine from r to $2r$; this can be seen also in Kyprianou [35].

Theorem 4.2.6 *If we define $\mathbb{Q}^x_\lambda := \tilde{\mathbb{Q}}^x_\lambda|_{\mathcal{F}_\infty}$, then $\mathbb{Q}^x_\lambda$ is a measure on $\mathcal{F}_\infty$ and*

$$\left.\frac{\mathrm{d}\mathbb{Q}^x_\lambda}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)},$$

so that under $\mathbb{Q}^x_\lambda$ (and therefore also under $\tilde{\mathbb{Q}}^x_\lambda$) the branching-diffusion point process $\mathbb{X}_t$ has exactly the pathwise construction given in definition 4.2.3.

There are at least two ways to prove this result: Kyprianou [35] bases his proof on a decomposition of the measure $\tilde{P}$ as a product of measures for the spine's motion, the fission-counting process n_t, and measures on the sub-trees born from the spine. Because of our using filtrations the way we do, we have an alternative.

Proof of Theorem 4.2.6: We have seen in te proof of Theorem 3.4.4 that the change of measure (4.7) projects onto the sub-algebra $\mathcal{F}_t$ as a conditional expectation:

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}^x_\lambda}{\mathrm{d}\tilde{P}^x}\right|_{\mathcal{F}_t} = e^{-\lambda x}\, \tilde{P}\big(e^{-rt}2^{n_t}e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t} | \mathcal{F}_t\big).$$

Bearing in mind that $2^{n_t} = \prod_{v<\xi_t} 2$, if we use the representation (3.4) we get

$$\begin{aligned}
\tilde{P}\big(e^{-rt}2^{n_t}e^{\lambda\xi_t-\frac{1}{2}\lambda^2 t}|\mathcal{F}_t\big) &= \tilde{P}\Big(e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t} \times \prod_{v<u} 2 \times 1_{(\xi_t=u)}|\mathcal{F}_t\Big) \\
&= e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t} \times \prod_{v<u} 2 \times \tilde{P}(\xi_t = u|\mathcal{F}_t) \\
&= e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t} \times \prod_{v<u} 2 \times \prod_{v<u}\frac{1}{2} \\
&= e^{-rt}\sum_{u\in N_t} e^{\lambda X_u(t)-\frac{1}{2}\lambda^2 t} = Z_\lambda(t).
\end{aligned}$$

Here, the result $\tilde{P}(\xi_t = u|\mathcal{F}_t) = \prod_{v<u}\frac{1}{2}$ is the case of (3.3) for binary-splitting. Thus

$$e^{-\lambda x}\,\tilde{P}\big(e^{-rt}2^{n_t}e^{\lambda\xi_t-\frac{1}{2}\lambda^2 t}|\mathcal{F}_t\big) = \frac{Z_\lambda(t)}{Z_\lambda(0)}.$$

□

Without loss of generality we throughout suppose that $\lambda \leq 0$, since the cases are symmetrical. The next theorem on $\mathcal{L}^p$-convergence of the martingale for $p > 1$ was originally proven in Neveu [44] by classical techniques, and clearly represents an extension to Kyprianou's result stated in Theorem 4.2.1. The second part is exactly the same result as stated in Kyprianou's theorem and therefore we do not repeat the proof, only giving a proof of the first part.

Theorem 4.2.7 *For each $x \in \mathbb{R}$, and for each $p \in [1, 2]$:*

1. *The martingale Z_λ is $\mathcal{L}^p(P^x)$-convergent if $p\lambda^2/2 < r$.*
2. *Almost surely under P^x, $Z_\lambda(\infty) = 0$ when $\lambda^2/2 \geq r$.*

We note that Neveu's result was actually based on a birth rate of $r = 1$, but the generalization to any $r > 0$ is trivial. Just before we proceed to the proof we recall a naturally occurring eigenvalue that will also appear in later models.

Definition 4.2.8 *For each $\lambda \in \mathbb{R}$ we define $E_\lambda := \frac{1}{2}\lambda^2 + r$.*

Then Z_λ can be re-written as

$$Z_\lambda(t) = \sum_{u\in N_t} e^{\lambda X_u(t) - E_\lambda t},$$

and simple algebra reveals

$$p\lambda^2/2 < r \quad \Leftrightarrow \quad pE_\lambda - E_{p\lambda} > 0.$$

Proof of part 1 of Theorem 4.2.7: We are going to prove that for every $p \in (1, 2]$ the martingale Z_λ is $\mathcal{L}^p(P)$-convergent if $pE_\lambda - E_{p\lambda} > 0$. Furthermore, since $P^x(Z_\lambda(t)) = e^{\lambda x}P^0(Z_\lambda(t))$ we do not lose generality if we suppose that $x = 0$, and from now on this is implicit and we do not use the superscript on the measures.

From the change of measure in Theorem 4.2.4 or Theorem 4.2.6 it is clear that

$$P(Z_\lambda(t)^p) = P(Z_\lambda(t)^{p-1} Z_\lambda(t)) = \mathbb{Q}_\lambda(Z_\lambda(t)^q),$$

where $q := p - 1$.

Theorem 4.2.9 *If $q \in [0,1]$ then $Z_\lambda(t)^q$ is a $\mathbb{Q}_\lambda$-submartingale.*

Proof: If $q \in [0,1]$ it follows from Jensen's inequality that $Z_\lambda(t)^{1+q}$ is a P-submartingale. This means that for $t \geq s$,

$$P\big(Z_\lambda(t)^{1+q}|\mathcal{F}_s\big) \geq Z_\lambda(s)^{1+q}, \qquad P\text{-a.s.}$$

or equivalently, for all $F \in \mathcal{F}_s$,

$$P\big(Z_\lambda(t)^{1+q}; F\big) \geq P\big(Z_\lambda(s)^{1+q}; F\big). \tag{4.8}$$

But this inequality is exactly the same as:

$$\text{for all } F \in \mathcal{F}_s, \quad \mathbb{Q}_\lambda\big(Z_\lambda(t)^q; F\big) \geq \mathbb{Q}_\lambda\big(Z_\lambda(s)^q; F\big),$$

and therefore we conclude that $Z_\lambda(t)^q$ is a $\mathbb{Q}_\lambda$-submartingale. □

Our aim is to prove that $\mathbb{Q}_\lambda(Z_\lambda(t)^q)$ is bounded for $t \in \mathbb{R}$, since then $Z_\lambda(t)$ must be bounded in $\mathcal{L}^p(P)$ and Doob's submartingale theorem will then imply that Z_λ is convergent in $\mathcal{L}^p(P)$.

As we have seen from Kyprianou's proof, the σ-algebra $\tilde{\mathcal{G}}_\infty$ gives us the very important *spine-decomposition* of the martingale Z_λ:

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big) = \sum_{k=1}^{n_t} e^{\lambda \xi_{S_k} - E_\lambda S_k} + e^{\lambda \xi_t - E_\lambda t},$$

where the sum is taken to equal 0 if $n_t = 0$. We continue with the conditional form of Jensen's inequality, which says that for $q \in (0,1]$:

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)^q|\tilde{\mathcal{G}}_\infty\big) \leq \tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big)^q. \tag{4.9}$$

The spine decomposition is a sum, and from Neveu's original proof we use the following simple inequality:

Proposition 4.2.10 *If $q \in (0,1]$ and $u, v > 0$ then $(u+v)^q \leq u^q + v^q$,*

to obtain,

$$\tilde{\mathbb{Q}}_\lambda\big(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty\big)^q \leq \sum_{k=1}^{n_t} e^{q\lambda \xi_{S_k} - qE_\lambda S_k} + e^{q\lambda \xi_t - qE_\lambda t}.$$

As written this spine decomposition is made up of two terms, and since they play a central role from here on we name them explicitly:

$$\textbf{spine term} := e^{q\lambda\xi_t - qE_\lambda t}, \quad \textbf{sum term} := \sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k}.$$

Taking $\tilde{\mathbb{Q}}_\lambda$-expectations of this spine decomposition and using (4.9),

$$\begin{aligned}\mathbb{Q}_\lambda(Z_\lambda(t)^q) = \tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)^q) &\leq \tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k}-qE_\lambda S_k} + e^{q\lambda\xi_t - qE_\lambda t}\Big) \\ &= \tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k}-qE_\lambda S_k}\Big) + \tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_t - qE_\lambda t}\Big), \end{aligned} \tag{4.10}$$

and the proof of $\mathcal{L}^p$-boundedness will be complete once we show that this RHS is bounded in t.

The spine term: Changing from $\tilde{P}$ to $\tilde{\mathbb{Q}}_\lambda$ gives the spine a drift of λ, and therefore the change-of-measure for just the spine's motion (i.e. on the algebra $\mathcal{G}_t$) is carried out by the martingale $e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}$:

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_t - qE_\lambda t}\Big) &:= \tilde{P}\Big(e^{q\lambda\xi_t - qE_\lambda t} \times e^{\lambda\xi_t - \frac{1}{2}\lambda^2 t}\Big), \\ &= e^{-(pE_\lambda - E_{p\lambda})t}\, \tilde{P}\Big(e^{p\lambda\xi_t - \frac{1}{2}(p\lambda)^2 t}\Big), \\ &= e^{-(pE_\lambda - E_{p\lambda})t}, \end{aligned} \tag{4.11}$$

since the second-line term $e^{p\lambda\xi_t - \frac{1}{2}(p\lambda)^2 t}$ is also a $\tilde{P}$-martingale, because under $\tilde{P}$ the process ξ_t is just a Brownian motion.

The sum term: Under the measure $\tilde{\mathbb{Q}}_\lambda$ we know that the fission times S_u on the spine occur as a Poisson process of rate $2r$. Appealing to standard results from Poisson theory (see [26] for example) we can therefore write the sum term as an integral:

$$\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k}-qE_\lambda S_k}\Big) = \tilde{\mathbb{Q}}_\lambda\Big(\int_0^t 2r\, e^{q\lambda\xi_s - qE_\lambda s}\,\mathrm{d}s\Big).$$

In this integral all the terms are positive and so Fubini's theorem can be used, which along with the equality (4.11) above gives

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k}-qE_\lambda S_k}\Big) &= \int_0^t 2r\,\tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_s - qE_\lambda s}\Big)\,\mathrm{d}s \\ &= 2r\int_0^t e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s \\ &= \frac{2r}{pE_\lambda - E_{p\lambda}}\Big[1 - e^{-(pE_\lambda - E_{p\lambda})t}\Big], \qquad \text{if } pE_\lambda \neq E_{p\lambda}. \end{aligned} \tag{4.12}$$

Thus we have found an explicit upper-bound:

$$P(Z_\lambda(t)^p) \leq \frac{2r}{pE_\lambda - E_{p\lambda}}\Big[1 - e^{-(pE_\lambda - E_{p\lambda})t}\Big] + e^{-(pE_\lambda - E_{p\lambda})t}. \tag{4.13}$$

If $pE_\lambda - E_{p\lambda} > 0$ this clearly implies that $P(Z_\lambda(t)^p)$ will remain bounded as $t \to \infty$, which together with Doob's theorem will complete the proof of the first part of Theorem 4.2.7. □

4.2.3 Random family sizes

In the binary-branching model of BBM that we just considered, at each fission time two particles are produced. Kyprianou [35] deals with a model in which at a fission time an individual u may split into a $1 + A$ particles where A is an integer-valued random variable chosen independently of the position $X_u(t)$ of the individual u, with general distribution:

$$P(A = i) = p_i, \qquad i \in \{0, 1, \ldots\},$$

giving an average size of $m := P(A) = \sum_{i=0}^{\infty} i\, p_i$. This introduction of random family sizes implies a small change in the form of the branching martingale:

Theorem 4.2.11 *For any $\lambda \in \mathbb{R}$,*

$$Z_\lambda(t) := \sum_{u \in N_t} e^{-rmt} e^{\lambda X_u(t) - \frac{1}{2}\lambda^2 t} = \sum_{u \in N_t} e^{\lambda X_u(t) - E_\lambda t}$$

is a P-martingale, where $E_\lambda := \frac{1}{2}\lambda^2 + rm$.

We remind the reader that the filtration $(\tilde{\mathcal{G}}_t)_{t \geq 0}$ also includes knowledge of the sizes of the families produced by all fissions on the spine:

$$\tilde{\mathcal{G}}_t = \sigma(\mathcal{G}_t, \mathrm{node}_s(\xi) : s \leq t, A_u : u < \xi_t).$$

The most significant alteration for the measure change is that the distribution of the family sizes produced by fissions on the spine (but not off it) is *size-biased*:

Theorem 4.2.12 *If we define the measure $\mathbb{Q}^x_\lambda$ via*

$$\left.\frac{\mathrm{d}\mathbb{Q}^x_\lambda}{\mathrm{d}P^x}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = e^{-\lambda x} Z_\lambda(t),$$

then it follows that under $\mathbb{Q}^x_\lambda$ the point process $\mathbb{X}_t$ evolves as follows:

- *starting from position x, the original ancestor diffuses according to a Brownian motion on $\mathbb{R}$ with drift λ;*
- *at an accelerated rate $(1 + m)r$ the particle undergoes fission producing $1 + \tilde{A}$ particles, where the distribution of $\tilde{A}$ is still independent of the spine's motion but is* size-biased*:*

$$\tilde{\mathbb{Q}}_\lambda(\tilde{A} = i) = \frac{(i+1)p_i}{m+1}, \qquad i \in \{0, 1, \ldots\}.$$

- *with equal probability, one of these offspring particles is selected;*
- *this chosen particle repeats stochastically the behaviour of the parent with the size-biased offspring distribution;*
- *the other particles initiate, from their birth position, an independent copy of a $P^{\cdot}$ branching Brownian motion with branching rate r and family-size distribution given by A (which is without the size-biasing).*

This size-biasing was noted in the Lyons *et al* papers [40, 34, 41] , and is common feature of the spine approach for branching models with random numbers of offspring.

As we have seen in chapter 3, the measures P and $\mathbb{Q}_\lambda$ on the $\mathcal{F}_\infty$ can be extended to $\tilde{P}$ and $\tilde{\mathbb{Q}}_\lambda$ on $\tilde{\mathcal{F}}_\infty$; or equivalently, we can define P and $\mathbb{Q}_\lambda$ as the projections onto $\mathcal{F}_\infty$ of the measures $\tilde{P}$ and $\tilde{\mathbb{Q}}_\lambda$ defined on $\tilde{\mathcal{F}}_\infty$.

The theorem on the $\mathcal{L}^p$-convergence of Z_λ is now slightly modified to take into account the random distribution of the family sizes:

Theorem 4.2.13 *For each $x \in \mathbb{R}$, and for each $p \in [1,2]$:*

1. *The martingale Z_λ is $\mathcal{L}^p(P^x)$-convergent if $p\lambda^2/2 < rm$ and $P(A^p) < \infty$.*
2. *Almost surely under P^x, $Z_\lambda(\infty) = 0$ when $\lambda^2/2 \geq rm$.*

The second part of this proof has been proven by Kyprianou [35], so we only here prove the first part of this theorem.

In fact the proof is not very different from the binary-splitting case since the spine decomposition is different only in the sum term:

$$\tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)|\tilde{\mathcal{G}}_\infty) = \sum_{k=1}^{n_t} A_u e^{\lambda\xi_{S_k} - E_\lambda S_k} + e^{\lambda\xi_t - E_\lambda t},$$

and therefore we go rather more quickly here. As before we can use Jensen's inequality and Proposition 4.2.10 to arrive at

$$\mathbb{Q}_\lambda(Z_\lambda(t)^q) = \tilde{\mathbb{Q}}_\lambda(Z_\lambda(t)^q) \leq \tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_u^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k} + e^{q\lambda\xi_t - qE_\lambda t}\Big).$$

The spine term can be dealt with as at (4.11), and the sum term can be written as an integral, but in order to deal with the random number of offspring we first use conditioning (without knowledge of the family sizes) to replace the term A_u^q with an expectation:

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_u^q e^{q\lambda\xi_{S_k} - qE_\lambda S_k}|\mathcal{G}_t\Big) &= \sum_{k=1}^{n_t} \tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) e^{q\lambda\xi_{S_k} - qE_\lambda S_k} \qquad \text{by independence,}\\ &= \tilde{\mathbb{Q}}_\lambda(\tilde{A}^q)\sum_{k=1}^{n_t} e^{q\lambda\xi_{S_k} - qE_\lambda S_k}. \qquad (4.14)\end{aligned}$$

The term $\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q)$ is guaranteed to be finite if $P(A^p)$ is:

Lemma 4.2.14 *If $P(A^p) < \infty$ then $\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) < \infty$.*

Proof:

$$\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q) = \sum_{i=1}^\infty i^q \frac{i+1}{m+1} p_i = \frac{1}{m+1}\Big(\sum_{i=1}^\infty i^p p_i + \sum_{i=1}^\infty i^q p_i\Big) = \frac{P(A^p) + P(A^q)}{m+1} < \frac{2P(A^p)}{m+1}.$$

Taking expectations of both sides of (4.14), converting the sum to an integral and then using Fubini's theorem gives:

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda\Big(\sum_{k=1}^{n_t} A_u^q e^{q\lambda\xi_{S_k}-qE_\lambda S_k}\Big) &= \tilde{\mathbb{Q}}_\lambda(\tilde{A}^q)\int_0^t 2r\,\tilde{\mathbb{Q}}_\lambda\Big(e^{q\lambda\xi_s-qE_\lambda s}\Big)\,\mathrm{d}s \\ &= 2r\,\tilde{\mathbb{Q}}_\lambda(\tilde{A}^q)\int_0^t e^{-(pE_\lambda-E_{p\lambda})s}\,\mathrm{d}s \qquad\qquad \text{by (4.11).}\end{aligned}$$

Thus the condition $P(A^p) < \infty$ in the first part of Theorem 4.2.13 means that $\mathbb{Q}_\lambda(Z_\lambda(t)^q)$ will be bounded if $pE_\lambda - E_{p\lambda} > 0$, and it can be shown that with $E_\lambda = \frac{1}{2}\lambda^2 + rm$,

$$pE_\lambda - E_{p\lambda} > 0 \quad \Leftrightarrow \quad p\lambda^2/2 < rm,$$

completing the proof.

□

4.3 The finite-type branching diffusion

We recall the branching diffusion from chapter 2. For a fixed $n \in \mathbb{N}$ we are given two sets of positive constants $a(1), \dots, a(n)$ and $r(1), \dots, r(n)$ and consider a typed branching diffusion in which the type of each particle moves as a finite, irreducible and time-reversible Markov chain on the set $I := \{1, \dots, n\}$ with Q-matrix θQ (θ is a strictly positive constant that could be considered as the *temperature* of the system) and invariant measure $\pi = \{\pi(1), \dots, \pi(n)\}$. The spatial movement of a particle of type y is a driftless Brownian motion with instantaneous variance $a(y)$, so that if $\big(X_u(t), Y_u(t)\big) \in \mathbb{R} \times I$ is the space-type location of individual u at time t then we have

$$\mathrm{d}X_u(t) = a(Y_u(t))\,\mathrm{d}B_t$$

for a Brownian motion B_t. Fission of a particle of type y occurs at a rate $r(y)$ to produce two particles at the same space-type location as the parent.

We define $J := \mathbb{R} \times I$, and suppose that the configuration of this whole branching diffusion at time t is given by the J-valued point process $\mathbb{X}_t = \big\{\big(X_u(t), Y_u(t)\big) : u \in N_t\big\}$ where N_t is the set of individuals alive at time t. Let the measures $\big\{P^{x,y} : (x, y) \in \mathbb{R}^2\big\}$ on the filtered space $(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0})$ be such that under $P^{x,y}$ the initial ancestor starts at (x, y) and $\mathbb{X}_t$ becomes the above-described branching diffusion – that is, we are supposing that $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration generated by the point process $\mathbb{X}_t$.

Using the ideas laid out in chapter 3 we can further extend these probability measures to get the probabilities $\big\{\tilde{P}^{x,y} : (x, y) \in J\big\}$ defined on $(\tilde{\mathcal{F}}_t)_{t\geq 0}$, the natural filtration with a spine, under which the branching diffusion $\mathbb{X}_t$ evolves according to the description given above, and where the spine is chosen uniformly so that the space-type location (ξ_t, η_t) of the spine behaves stochastically like any one of the single particles in the branching diffusion – thus η_t is a Markov chain on I with Q-matrix θQ and ξ_t is a diffusion on $\mathbb{R}$ satisfying $\mathrm{d}\xi_t = a(\eta_t)\,\mathrm{d}\tilde{B}_t$ for a $\tilde{P}$-Brownian motion $\tilde{B}_t$.

We have already given a classical proof of the following theorem, and in this section will give a spine-based proof.

Theorem 4.3.1 *Suppose that $\lambda \leq 0$.*

1. *For every $p \in (1,2]$ the martingale Z_λ is $\mathcal{L}^p$-bounded provided that $pE_\lambda - E_{p\lambda} > 0$. In fact this inequality holds for some $p \in (1,2]$ if and only if $\lambda \in (\tilde{\lambda}(\theta), 0]$ (see Corollary 2.4.13).*
2. *Almost surely (under P), $Z_\lambda(\infty) = 0$ if $\lambda < \tilde{\lambda}(\theta)$.*

The question of what happens at the critical $\lambda = \tilde{\lambda}(\theta)$ is not considered, but based on the work by Harris [23] or Kyprianou [35] for BBM, we conjecture (but do not prove) that the martingale limit $Z_\lambda(\infty)$ is null for λ at the critical value.

Our spine approach means that we should like to make a change of measure with $Z_\lambda(t)$ as the Radon-Nikodym derivative, as we did for BBM at Theorem 4.2.6. Here a pathwise construction could be made analogous to the one laid out for BBM in Definition 4.2.3, but we have seen the better alternative:

Definition 4.3.2 *For each $\lambda \leq 0$ we define a measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\tilde{\mathcal{F}}_t} := \frac{1}{v_\lambda(y)e^{\lambda x}}\, e^{-\int_0^t R(\eta_s)\,\mathrm{d}s} 2^{n_t} \times v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda \xi_t - E_\lambda t}. \tag{4.15}$$

This Radon-Nikodym derivative is going to introduce drift to the spine via the martingale term $v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}$, as we see in the following section. The other martingale term $e^{-\int_0^t R(\eta_s)\,\mathrm{d}s} 2^{n_t}$ is a martingale for the Cox process n_t that will increase the rate at which fission times occur on the spine – see Kyprianou [35] for more details; the terms $v_\lambda(y)^{-1}e^{-\lambda x}$ are just normalizing constants. A proof like that for Theorem 4.2.6 with the conditional expectation works here to give:

Theorem 4.3.3 *If we define $\mathbb{Q}_\lambda^{x,y} := \tilde{\mathbb{Q}}_\lambda^{x,y}|_{\mathcal{F}_\infty}$, then $\mathbb{Q}_\lambda^{x,y}$ is a measure on $\mathcal{F}_\infty$ and*

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \frac{Z_\lambda(t)}{v_\lambda(y)e^{\lambda x}}.$$

4.3.1 The spine process (ξ_t, η_t) under $\tilde{\mathbb{Q}}_\lambda$

In the BBM model it was clear to see that the spine process ξ_t received a drift under the measure $\tilde{\mathbb{Q}}$; something similar happens here:

Lemma 4.3.4 *Under $\tilde{\mathbb{Q}}_\lambda$ the spine process (ξ_t, η_t) has generator:*

$$\mathcal{H}_\lambda F(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + a(y)\lambda\frac{\partial F}{\partial x} + \sum_{j\in I} \theta Q_\lambda(y,j)F(x,j), \tag{4.16}$$

where Q_λ is an honest Q-matrix:

$$\theta Q_\lambda(i,j) = \begin{cases} \theta Q(i,j)\frac{v_\lambda(j)}{v_\lambda(i)} & \text{if } i \neq j \\ \theta Q(i,i) + \frac{\lambda^2}{2}a(i) - E_\lambda + r(i) & \text{if } i = j \end{cases}$$

with invariant measure $\pi_\lambda = v_\lambda^2 \pi$.

Thus under $\tilde{\mathbb{Q}}_\lambda$ the Q-matrix (generator) of η_t is changed, and the process $\xi_t \in \mathbb{R}$ is given an instantaneous drift of $a(\eta_t)\lambda$. The form of this above generator can be obtained from the theory of Doob's h-transforms, due to the fact that on the algebra $\mathcal{G}_t$ the change of measure is given by:

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{G}_t} = \frac{1}{v_\lambda(y)e^{\lambda x}} v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}. \tag{4.17}$$

which as we have noted is exactly the same as the single-particle change of measure (2.18).

The long-term behaviour under $\tilde{\mathbb{Q}}_\lambda$ of the spine diffusion ξ_t can be retrieved from the generator (4.16) and the form for the derivative of E_λ stated in Theorem 2.4.9:

Theorem 4.3.5 *Under $\tilde{\mathbb{Q}}$ the long-term drift of the spine diffusion ξ_t is*

$$\lim_{t\to\infty} t^{-1}\xi_t = E'_\lambda.$$

Proof: From the generator stated at (4.16) we can write:

$$\xi_t = B\left(\int_0^t a(\eta_s)\mathrm{d}s\right) + \lambda\int_0^t a(\eta_s)\mathrm{d}s,$$

where $B(t)$ is a $\tilde{\mathbb{Q}}_\lambda$-Brownian motion. Then by the ergodic theorem and the fact that $\pi_\lambda = v_\lambda^2\pi$:

$$t^{-1}\xi_t \to \lambda\sum_{y\in I} a(y)\pi_\lambda(y) = \lambda\sum_{y\in I} a(y)v_\lambda^2(y)\pi(y) = \lambda\langle Av_\lambda, v_\lambda\rangle_\pi = E'_\lambda.$$

□

Direct calculation from the definition of c_λ given in definition 2.4.10 gives $E'_\lambda = -c_\lambda - \lambda c'_\lambda$, and therefore $t^{-1}(\xi_t + c_\lambda t) \to -\lambda c'_\lambda$, whence we have

Corollary 4.3.6 *The critical value $\tilde{\lambda}(\theta)$ represents a crucial difference in the long-term behaviour of $\xi_t + c_\lambda t$:*

- *$\xi_t + c_\lambda t$ drifts off to $+\infty$ if $\lambda \in (\tilde{\lambda}(\theta), 0]$,*
- *$\xi_t + c_\lambda t$ drifts off to $-\infty$ if $\lambda < \tilde{\lambda}(\theta)$.*

This second fact will be important in showing that the martingale limit is null when $\lambda < \tilde{\lambda}(\theta)$.

Before moving on to our proof of Theorem 4.3.1, and just to drive home the point we state the pathwise construction of $\tilde{\mathbb{Q}}_\lambda^x$:

- starting from position x, the original ancestor diffuses according to the generator (4.16);
- at rate $2R(\eta_t)$ the particle undergoes fission producing two particles;
- with equal probability, one of these two particles is selected to form the next node of the spine;
- this chosen particle repeats stochastically the behaviour of the parent;
- the other particle initiates, from its birth position, an independent copy of a $P^\cdot$ branching Brownian motion with branching rate $R(\cdot)$.

4.3.2 Proof of Theorem 4.3.1

Proof of Part 1 of Theorem 4.3.1: Suppose $p \in (1,2]$, then with $q := p-1$ a slight modification of the BBM proof arrives at

$$\begin{aligned}P^{x,y}(Z_\lambda(t)^p) &= \mathbb{Q}_\lambda^{x,y}(Z_\lambda(t)^q) = \tilde{\mathbb{Q}}_\lambda^{x,y}(Z_\lambda(t)^q)\\ &\le \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u}-qE_\lambda S_u}\Big) + \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big)\end{aligned}$$

and the proof of $\mathcal{L}^p$-boundedness will be complete once we show that this RHS expectation is bounded in t.

The spine term. It is always useful to first focus on the spine term, since we can change the measure with (4.17) to get

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t-qE_\lambda t}\Big) &= \tilde{P}^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t-qE_\lambda t} \times \frac{v_\lambda(\eta_t)\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}}{v_\lambda(y)e^{\lambda x}}\Big)\\ &= e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times e^{-(pE_\lambda - E_{p\lambda})t}\, \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big). \qquad (4.18)\end{aligned}$$

Bearing in mind that η_t is a finite-state irreducible Markov chain and therefore ergodic, and given $\pi_\lambda(y) = v_\lambda(y)^2\pi(y)$, the following result is immediate,

Lemma 4.3.7 *In the finite-type model, for any $\lambda, p \in \mathbb{R}$, the expectation*

$$\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big)$$

is positive, bounded and convergent with

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big) = \frac{v_\lambda^p}{v_{p\lambda}}\cdot \pi_{p\lambda} = \langle v_\lambda^p, v_{p\lambda}\rangle_\pi,$$

With boundedness and convergence of $\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\big)$ ensured by the above lemma, it follows from (4.18) that in the long term the growth or decay of the spine term is determined by the term $e^{-(pE_\lambda - E_{p\lambda})t}$.

The sum term. We know that under $\tilde{\mathbb{Q}}_\lambda$ the fission times S_u on the spine occur as a *Cox* process – that is, conditional on knowing η, the times occur as a Poisson process of rate $2R(\eta_s)$. Therefore, if we condition on $\mathcal{G}_t$ which knows about η_s at all times $0 \le s \le t$ we can transform the sum into an integral and use Fubini's theorem:

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u}-qE_\lambda S_u}\Big) &= \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u}-qE_\lambda S_u}\big|\mathcal{G}_t\Big)\Big)\\ &= \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\int_0^t 2R(\eta_s)\, v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s}\,\mathrm{d}s\Big)\\ &= \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(2R(\eta_s)v_\lambda(\eta_s)^q e^{q\lambda\xi_s-qE_\lambda s}\Big)\,\mathrm{d}s. \qquad (4.19)\end{aligned}$$

The change of measure used in (4.18) can be used on the sum term in its integral form,

$$\int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(2R(\eta_s)v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s}\Big)\,\mathrm{d}s = e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)}\int_0^t \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big)e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s$$
$$\leq K(p,\lambda,y)\int_0^t e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s$$

where

$$K(p,\lambda,x,y) := e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \sup_{s\geq 0}\tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(2R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big), \tag{4.20}$$

is finite by a simple adaptation of the above lemma since the birth rates $R(\cdot)$ are clearly bounded.

Having dealt with the spine term and the sum term, we have therefore obtained an explicit upper-bound, (if $pE_\lambda - E_{p\lambda} \neq 0$)

$$P^{x,y}\big(Z_\lambda(t)^p\big) \leq K(p,\lambda,x,y)\Big\{\frac{2\bar{r}}{pE_\lambda - E_{p\lambda}}\Big[1 - e^{-(pE_\lambda - E_{p\lambda})t}\Big] + e^{-(pE_\lambda - E_{p\lambda})t}\Big\},$$

and hence:

$$pE_\lambda - E_{p\lambda} > 0 \quad \Rightarrow \quad P^{x,y}\big(Z_\lambda(t)^p\big) \text{ is bounded for all } t.$$

Together with the facts laid out in Lemma 2.4.13 this completes the proof of the first part of Theorem 4.3.1.

□

Proof of Part 2 of Theorem 4.3.1: For the second part we again use the spine term as a lower bound to $Z_\lambda(t)$:

$$Z_\lambda(t) \geq v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t} = v_\lambda(\eta_t)e^{\lambda(\xi_t + c_\lambda t)}.$$

We aim to show that under $\tilde{\mathbb{Q}}_\lambda$ this spine term is almost-surely unbounded whenever $\lambda < \tilde{\lambda}(\theta)$, leading to the result

$$\limsup_{t\to\infty} Z_\lambda(t) = \infty, \qquad \mathbb{Q}_\lambda\text{-a.s.}$$

which with Theorem 4.2.2 will imply

$$Z_\lambda(\infty) = 0 \qquad P\text{-a.s.}$$

It was seen that as a consequence of Corollary 4.3.5, $\xi_t + c_\lambda t \to -\infty$ almost surely under $\tilde{\mathbb{Q}}_\lambda$ whenever $\lambda < \tilde{\lambda}(\theta)$, and therefore

$$\limsup_{t\to\infty} v_\lambda(\eta_t)e^{\lambda(\xi_t + c_\lambda t)} = \infty$$

because $v(\eta_t) \geq \min_{y\in I} v_\lambda(y) > 0$. This concludes the proof of part 2.

□

4.3.3 Random family sizes

As we did for the BBM model earlier, we now consider extending our typed branching diffusion to allow a type-dependent randomness in the number of offspring that each individual produces of its own type when it undergoes fission. Thus we suppose that under the probabilities $\{\tilde{P}^{x,y} : (x,y) \in J\}$ defined on $(\tilde{\mathcal{F}}_t)_{t\geq 0}$, if a type-$w$ particle undergoes fission then the number $1 + A_u$ of type-w offspring it produces is distributed like a random variable $1 + A(w) \in \{1, 2, \ldots\}$ with

$$P\big(A(w) = i\big) = p_i(w), \qquad i \in \{0, 1, \ldots\},$$

and mean $P\big(A(w)\big) = m(w)$. For the reasons seen in Lemma 4.2.14, we suppose that there is some $p > 1$ such that for each $w \in I$,

$$P\big(A(w)^p\big) < \infty.$$

In the spine composition below we shall need to refer to the tilted distribution $\tilde{\mathbb{Q}}_\lambda(\tilde{A}(w)^q)$, and from Lemma 4.2.14 we know what this value is:

Definition 4.3.8 *For $p \in (1, 2]$ and $q = p - 1$, we define*

$$M_q(w) := \frac{P(A^p(w)) + P(A^q(w))}{m(w) + 1}.$$

The form of the martingale is unchanged by the random offspring numbers, but the relationship between v_λ and E_λ is altered to account for the average family size being $m(w)$ when a fission occurs for a type-w particle:

Definition 4.3.9

$$Z_\lambda(t) := \sum_{k=1}^{N(t)} v_\lambda(Y_k(t))\, e^{\lambda X_k(t) - E_\lambda t}, \tag{4.21}$$

where v_λ is a strictly-positive vector on I normalized so that $\langle v_\lambda, v_\lambda \rangle_\pi = 1$, and $E_\lambda \in \mathbb{R}$ satisfying

$$\Big(\frac{1}{2}\lambda^2 A + \theta Q + mR\Big) v_\lambda = E_\lambda v_\lambda,$$

in which mR is the diagonal matrix

$$mR = \mathrm{diag}[m(1)R(1), m(2)R(2), \ldots, m(n)R(n)].$$

We have seen in three examples that for the change of measure (on sub-filtration $(\mathcal{F}_t)_{t\geq 0}$) to have $Z_\lambda(t)$ as its Radon-Nikodym derivative, we should expect the new measure to introduce three changes:

- it should *affect the motion* of the spine (ξ_t, η_t) in some way;
- it should *increase the rate* at which the fissions occur on the spine;
- it should cause the distributions of families produced from the spine to be *size-biased.*

These three features will be brought about by three martingales.

Lemma 4.3.10

$$v_\lambda(\eta_t)\, e^{\int_0^t mR(\eta_s)\,\mathrm{d}s} e^{\lambda\xi_t - E_\lambda t}$$

is a P-martingale that will introduce a drift to the spine's spatial motion ξ_t and will change the Q-matrix (generator) of the spine's type motion η_t.

Lemma 4.3.11

$$e^{-\int_0^t mR(\eta_s)\,\mathrm{d}s} \prod_{v<\xi_t} (1+m(\eta_{S_v}))$$

is a P-martingale that will increase the rate at which fission times occur on the spine from $R(\eta_t)$ to $(1+m(\eta_t))R(\eta_t)$.

Lemma 4.3.12

$$\prod_{v<\xi_t} \frac{1+A_v}{1+m(\eta_{S_v})}$$

is a P-martingale that will cause the family distribution on the spine to be size-biased, so that under the new measure

$$Prob(A=i) = \frac{(1+i)p_i(\eta_{S_v})}{1+m(\eta_{S_v})}, \qquad i \in \{0,1,\ldots\}.$$

Therefore we change the measure $\tilde{P}$ by the product of these three martingales – for which some of the terms cancel:

Definition 4.3.13 *For each $\lambda \leq 0$ we define a measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\tilde{\mathcal{F}}_t} := \frac{1}{v_\lambda(y)e^{\lambda x}} \prod_{v<\xi_t}(1+A_v) \times v_\lambda(\eta_t)e^{\lambda\xi_t - E_\lambda t}. \tag{4.22}$$

Once again, a proof using the conditional expectation of this measure-change martingale confirms:

Theorem 4.3.14 *If we define $\mathbb{Q}_\lambda^{x,y} := \tilde{\mathbb{Q}}_\lambda^{x,y}|_{\mathcal{F}_\infty}$, then $\mathbb{Q}_\lambda^{x,y}$ is a measure on $\mathcal{F}_\infty$ and*

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \frac{Z_\lambda(t)}{v_\lambda(y)e^{\lambda x}}.$$

The geometry of E_λ is not significantly changed by the introduction of random offspring numbers, although $\tilde{\lambda}(\theta)$ will now be different. In fact, as far as E_λ is concerned, the introduction of random offspring numbers acts just like a multiplying term on the birth rates $R(y)$.

Theorem 4.3.15 *Suppose that $\lambda \leq 0$.*

1. *For every $p \in (1,2]$ the martingale Z_λ is $\mathcal{L}^p$-bounded provided that $pE_\lambda - E_{p\lambda} > 0$.*

2. *Almost surely (under P), $Z_\lambda(\infty) = 0$ if $\lambda < \tilde{\lambda}(\theta)$.*

Proof of part 1: We quickly cover the main points since they are very similar to the previous proof for the binary-splitting case. The spine decomposition leads us to:

$$P^{x,y}(Z_\lambda(t)^p) \le \mathbb{Q}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) + \mathbb{Q}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big),$$

By conditioning on which nodes make up the spine ξ and on the spine's motion (information in $\mathcal{G}_t$), we can replace the term A_u^q by its expectation:

$$\begin{aligned}\mathbb{Q}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\big|\xi, \mathcal{G}_\infty\Big) &= \sum_{u<\xi_t} \mathbb{Q}_\lambda^{x,y}(A_u^q) v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u} \\ &= \sum_{u<\xi_t} M_q(\eta_{S_u}) v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}.\end{aligned}$$

Taking expectations again,

$$\mathbb{Q}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) = \mathbb{Q}_\lambda^{x,y}\Big(\sum_{u<\xi_t} M_q(\eta_{S_u}) v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big).$$

The remainder of the proof now goes through as before when we write this sum as an integral, use Fubini's theorem and change the measure like at (4.18):

$$\mathbb{Q}_\lambda^{x,y}\Big(\sum_{u<\xi_t} A_u^q v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) \le \Big(\max_{i\in I} M_q(i)\Big)\ K(p,\lambda,x,y) \int_0^t e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s,$$

where

$$K(p,\lambda,x,y) := e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \sup_{s\ge 0} \mathbb{Q}_{p\lambda}^{0,y}\Big(2R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big) < \infty$$

as at (4.20). Thus if $p > 1$ is such that $P(A(w)^p) < \infty$ for all $w \in I$ we know that $M_q(w) < \infty$ also, whence $\max_{i\in I} M_q(i) < \infty$, and the growth of the sum term is once again dependent on the term $e^{-(pE_\lambda - E_{p\lambda})t}$, completing the proof of part 1. □

Proof of part 2: The fact that the geometry interprets random offspring numbers much as it would handle an increase in the birth rates is plain to see in the way that the Q-matrix Q_λ is defined:

Lemma 4.3.16 *Under $\tilde{\mathbb{Q}}_\lambda$ the spine process (ξ_t, η_t) has generator:*

$$\mathcal{H}_\lambda F(x,y) = \frac{1}{2}a(y)\frac{\partial^2 F}{\partial x^2} + a(y)\lambda\frac{\partial F}{\partial x} + \sum_{j\in I}\theta Q_\lambda(y,j)F(x,j), \tag{4.23}$$

where Q_λ is an honest Q-matrix:

$$\theta Q_\lambda(i,j) = \begin{cases} \theta Q(i,j)\frac{v_\lambda(j)}{v_\lambda(i)} & \text{if } i \ne j \\ \theta Q(i,i) + \frac{\lambda^2}{2}a(i) - E_\lambda + m(i)r(i) & \text{if } i = j \end{cases}$$

with invariant measure $\pi_\lambda = v_\lambda^2\pi$.

The proof for the binary-splitting model easily adapts:

$$\limsup_{t\to\infty} Z_\lambda(t) > \limsup_{t\to\infty} v_\lambda(\eta_t)e^{\lambda(\xi_t + c_\lambda t)} = \infty, \quad \text{if } \lambda < \tilde{\lambda}(\theta).$$

□

4.4 A continuous-type branching diffusion

The preceding finite-type model was originally inspired by the model that we now turn to, originally laid out in Harris and Williams [21]. In this model the type moves on the real line as an Orstein-Uhlenbeck process associated with the generator

$$Q_\theta := \frac{\theta}{2}\Big(\frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y}\Big), \qquad \text{with } \theta > 0 \text{ considered as the } \textit{temperature},$$

which has the standard normal density as its invariant distribution:

$$\phi(y) := (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}y^2}.$$

The spatial movement of a particle of type y is a driftless Brownian motion with instantaneous variance

$$A(y) := ay^2, \qquad \text{for some fixed } a > 0,$$

and fission of a particle of type y occurs at a rate

$$R(y) := ry^2 + \rho, \qquad \text{where } r, \rho > 0 \text{ are fixed},$$

to produce two particles at the same type-space location as the parent (we consider only binary splitting). The model has very different behaviour for low temperature values (i.e. low θ), but most studies have considered the high temperature regime where $\theta > 8r$. Also, the parameter λ must be restricted to an interval $(\lambda_{\min}, 0)$ in order for some of the model's parameters to remain in $\mathbb{R}$, where

$$\lambda_{\min} := -\sqrt{\frac{\theta - 8r}{4a}}.$$

Generally, *unboundedness* in a model's rates is a serious obstacle to classical proofs since they often depend on the expectation semigroup of the branching process, and unbounded rates tend to lead to unbounded *eigenfunctions*. Here this is the case, but the existence of a spectral theory for their particular expectation operator allowed Harris and Williams to get a sufficiently good bound in particular for a non-linear term (see Theorem 5.1 of [21]), and therefore to prove $\mathcal{L}^p$-convergence of the martingale.

We use the same notation as previously $\mathbb{X}_t = \{(X_u(t), Y_u(t)) : u \in N_t\}$ to denote the point process of space-type locations in $\mathbb{R} \times \mathbb{R}$, and suppose that the measures $\{\tilde{P}^{x,y} : (x,y) \in \mathbb{R}^2\}$ on the natural filtration with a spine $(\tilde{\mathcal{F}}_t)_{t\geq 0}$ are such that the initial ancestor starts at (x,y) and $(\mathbb{X}_t, (\xi_t, \eta_t))$ becomes the above-described branching diffusion with a spine.

4.4.1 The measure change

Although there are some significant differences, this model is similar in flavour to our finite-type model – it was in fact the inspiration for that finite model. There is a strictly-positive martingale Z_λ^- defined as

$$Z_\lambda^-(t) := \sum_{u \in N_t} v_\lambda(Y_u(t)) e^{\lambda X_u(t) - E_\lambda^- t}$$

where v_λ and E^-_λ are the eigenvector and eigenvalue associated with the operator:

$$Q_\theta + \frac{1}{2}\lambda^2 A(y) + R(y),$$

which is self-adjoint with respect to the inner-product defined as

$$\langle f, g\rangle_\phi := \int_{-\infty}^{\infty} f(y)g(y)\phi(y)\,\mathrm{d}y.$$

The eigenfunction v_λ is in $\mathcal{L}^2(\phi)$ and can be found explicitly as

$$v_\lambda(y) = K_\lambda e^{\psi^-_\lambda y^2}$$

where

$$\psi^-_\lambda := \frac{1}{4} - \frac{\mu_\lambda}{2\theta}, \quad \mu_\lambda := \frac{1}{2}\sqrt{\theta^2 - \theta(8r + 4a\lambda^2)},$$

are both positive for all $\lambda \in (\lambda_{\min}, 0)$, and $K_\alpha > 0$ is just a normalizing factor: $K_\lambda = \|e^{\psi^-_\lambda y^2}\|^{-1}_\phi$, which guarantees that

$$\|v_\lambda\|_\phi = 1.$$

The eigenvalue E^-_λ is given by

$$E^-_\lambda = \rho + \theta\psi^-_\lambda = \rho + \theta\Big(\frac{1}{4} - \frac{\mu_\lambda}{2\theta}\Big), \tag{4.24}$$

and it has been shown by Harris and Williams [21] that just as was the case in the finite-type model at Theorem 2.4.9,

$$\frac{\mathrm{d}E^-_\lambda}{\mathrm{d}\lambda} = \lambda\langle A(y), v^2_\lambda\rangle_\phi = \lambda\langle ay^2, v^2_\lambda\rangle_\phi = \frac{\lambda a\theta}{2\mu_\lambda}.$$

We again define the speed function $c^-_\lambda := -E^-_\lambda/\lambda$, and $\tilde{\lambda}(\theta) < 0$ is the unique point (on the negative axis) at which c^-_λ hits its minimum – further details are given in Harris and Williams [21]. Another important parameter is $\psi^+_\lambda := \frac{1}{4} + \frac{\mu_\lambda}{2\theta}$.

As for the finite-type model, we can prove

Lemma 4.4.1 *On $(\lambda_{\min}, 0)$ the function c^-_λ has just one minimum at the single point $\tilde{\lambda}(\theta)$, strictly increasing to $+\infty$ as $\lambda \uparrow 0$; we define $\tilde{c}(\theta) := c_{\tilde{\lambda}(\theta)}$ as this minimum value.*

Furthermore, for each $\lambda \in (\tilde{\lambda}(\theta), 0]$ there is some $p > 1$ such that $pE_\lambda - E_{p\lambda} > 0$; on the other hand, if $\lambda < \tilde{\lambda}(\theta)$ there is no such $p > 1$.

We are going to use spines to prove the result first given by Harris and Williams [21] using a classical-style proof similar to the one we gave in section 2.5 for the finite-type model.

Theorem 4.4.2 *Suppose that $\lambda \in (\lambda_{\min}, 0)$.*

1. *Let $p \in (1, 2]$. The martingale Z_λ is $\mathcal{L}^p$-bounded if both $pE_\lambda - E_{p\lambda} > 0$ and $p\psi^-_\lambda < \psi^+_{p\lambda}$.*
2. *Almost surely (under P), $Z_\lambda(\infty) = 0$ if $\lambda < \tilde{\lambda}(\theta)$.*

We note that the appearance of the extra condition $p\psi_\lambda^- < \psi_{p\lambda}^+$ will be explained as relating to the fact that, differently from the finite-type model, here we have to deal with the unboundedness that can result from having an continuous type space.

Once again, for each $\lambda \leq 0$ we define a measure $\tilde{\mathbb{Q}}_\lambda^{x,y}$ on $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\infty)$ via

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\tilde{\mathcal{F}}_t} := \frac{1}{v_\lambda(y)e^{\lambda x}} 2^{n_t} v_\lambda(\eta_t) e^{\lambda\xi_t - E_\lambda^- t}, \tag{4.25}$$

so that with $\mathbb{Q}_\lambda := \tilde{\mathbb{Q}}_\lambda|_{\mathcal{F}_\infty}$ we have

$$\left.\frac{\mathrm{d}\mathbb{Q}_\lambda^{x,y}}{\mathrm{d}P^{x,y}}\right|_{\mathcal{F}_t} = \frac{Z_\lambda(t)}{Z_\lambda(0)} = \frac{Z_\lambda(t)}{v_\lambda(y)e^{\lambda x}}.$$

The similarities between this continuous-type model and the finite-type model that it inspired continue, and under $\tilde{\mathbb{Q}}_\lambda$:

- the spine diffusion ξ_t has instantaneous drift $a\eta_t^2\lambda$;
- the type process η_t has generator $\frac{\theta}{2}\left(\frac{\partial^2}{\partial y^2} - \frac{2\mu_\lambda}{\theta} y \frac{\partial}{\partial y}\right)$ and invariant measure with density $\phi_\lambda(y)$ proportional to $v_\lambda^2(y)\phi(y)$, which is to say that it is a $N(0, \frac{\theta}{2\mu_\lambda})$ distribution;
- fission times on the spine occur at the accelerated rate of $2R(\eta_t)$;
- all particles not in the spine behave as if under the original measure P.

4.4.2 Proof of Theorem 4.4.2

Proof of Part 1: Suppose $p \in (1, 2]$. Then using the spine decomposition with Jensen's inequality and Proposition 4.2.10 we find,

$$P^{x,y}(Z_\lambda^-(t)^p) \leq \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) + \mathbb{Q}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big).$$

The spine term. On the algebra $\mathcal{G}_t$ the change of measure takes the form

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_\lambda^{x,y}}{\mathrm{d}\tilde{P}^{x,y}}\right|_{\mathcal{G}_t} = v_\lambda(y)^{-1} e^{-\lambda x}\, e^{\int_0^t R(\eta_s)\,\mathrm{d}s} v_\lambda(\eta_s) e^{\lambda\xi_t - E_\lambda^- t},$$

which we can use on the spine term to arrive at

$$\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(v_\lambda(\eta_t)^q e^{q\lambda\xi_t - qE_\lambda t}\Big) = e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \tilde{\mathbb{Q}}_{p\lambda}^{0,y}\Big(\frac{v_\lambda^p}{v_{p\lambda}}(\eta_t)\Big) e^{-(pE_\lambda - E_{p\lambda})t}. \tag{4.26}$$

The sum term. As for the finite-type model the fission times S_u on the spine occur as a Cox process and therefore

$$\begin{aligned}\tilde{\mathbb{Q}}_\lambda^{x,y}\Big(\sum_{u<\xi_t} v_\lambda(\eta_{S_u})^q e^{q\lambda\xi_{S_u} - qE_\lambda S_u}\Big) &= \int_0^t \tilde{\mathbb{Q}}_\lambda^{x,y}\Big(2R(\eta_s)\, v_\lambda(\eta_s)^q e^{q\lambda\xi_s - qE_\lambda s}\Big)\,\mathrm{d}s \\ &= e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \int_0^t \mathbb{Q}_{p\lambda}^{0,y}\Big(R(\eta_s)\,\frac{v_\lambda^p}{v_{p\lambda}}(\eta_s)\Big) e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s.\end{aligned}$$

Bringing together the results for the sum and spine terms, we have an upper bound

$$P^{x,y}(Z_\lambda(t)^p) \le e^{q\lambda x}\frac{v_{p\lambda}(y)}{v_\lambda(y)} \times \left\{ \tilde{\mathbb{Q}}^{0,y}_{p\lambda}\Big(\frac{v^p_\lambda}{v_{p\lambda}}(\eta_t)\Big)e^{-(pE_\lambda - E_{p\lambda})t} + \int_0^t \mathbb{Q}^{0,y}_{p\lambda}\Big(R(\eta_s)\,\frac{v^p_\lambda}{v_{p\lambda}}(\eta_s)\Big)e^{-(pE_\lambda - E_{p\lambda})s}\,\mathrm{d}s\right\}. \quad (4.27)$$

Once again the role of $pE_\lambda - E_{p\lambda} > 0$ is clear, but the new condition $p\psi^-_\lambda < \psi^+_{p\lambda}$ in part 1 of Theorem 4.4.2 is due to the term $\mathbb{Q}^{0,y}_{p\lambda}\big(R(\eta_s)\,\frac{v^p_\lambda}{v_{p\lambda}}(\eta_s)\big)$, which could potentially be unbounded.

Harris and Williams have shown in [21] that under $\tilde{\mathbb{Q}}^{0,y}_{p\lambda}$, the random variable η_s has the normal distribution

$$\mathrm{N}\Big(e^{-\mu_{p\lambda}s}y, \frac{\theta(1-e^{-2\mu_{p\lambda}s})}{2\mu_{p\lambda}}\Big) =: \mathrm{N}(\alpha, \beta)$$

and from the known form of the eigenfunctions we have the explicit result:

$$\tilde{\mathbb{Q}}^{0,y}_{p\lambda}\Big(R(\eta_s)\frac{v^p_\lambda}{v_{p\lambda}}(\eta_s)\Big) = \frac{K^p_\lambda}{K_{p\lambda}}\frac{1}{\sqrt{2\pi\alpha^2}}\int_{-\infty}^{\infty}(ru^2+\rho)e^{(p\psi^-_\lambda - \psi^-_{p\lambda})u^2}e^{-(u^2-\alpha)/2\beta}\,\mathrm{d}u. \quad (4.28)$$

which will be finite and bounded for all $s \ge 0$ if $p\psi^-_\lambda - \psi^-_{p\lambda} - \frac{\mu_{p\lambda}}{\theta} < 0$; here the term $\frac{\mu_{p\lambda}}{\theta}$ comes from the term $(2\beta)^{-1}$ in the distribution of η_s. But

$$p\psi^-_\lambda - \psi^-_{p\lambda} - \frac{\mu_{p\lambda}}{\theta} = p\psi^-_\lambda - \psi^+_{p\lambda}$$

and therefore the condition $p\psi^-_\lambda < \psi^+_{p\lambda}$ ensures that the expectation (4.28) stays bounded for all $s \ge 0$. Thus we deduce that with $p \in (1,2]$ and $\lambda \in (\lambda_{\min}, 0)$,

$$pE_\lambda - E_{p\lambda} < 0 \text{ and } p\psi^-_\lambda < \psi^+_{p\lambda} \quad \text{implies} \quad P^{x,y}(Z_\lambda(t)^p) \text{ is bounded for all } t \ge 0,$$

□

Proof of Part 2: The proof that we have seen in the finite-type model will work here unchanged, since under $\tilde{\mathbb{Q}}_\lambda$ the spatial motion is

$$\xi_t = B\left(\int_0^t a(\eta_s)\mathrm{d}s\right) + \lambda a\int_0^t \eta_s^2\mathrm{d}s,$$

and the type process η_s has invariant distribution $N(0, \frac{\theta}{2\mu_\lambda})$, whence

$$t^{-1}\xi_t \to \lambda a\theta/2\mu_\lambda = E'_\lambda.$$

From the fact that $E'_\lambda = -c_\lambda - \lambda c'_\lambda$, and Lemma 4.4.1 this implies that

$$t^{-1}(\xi_t + c_\lambda t) \to -\lambda c'_\lambda > 0$$

and therefore under $\tilde{\mathbb{Q}}_\lambda$ the diffusion $\xi_t + c_\lambda t$ drifts off to $-\infty$ if $\lambda < \tilde{\lambda}(\theta)$ whence

$$\tilde{\mathbb{Q}}^{x,y}_\lambda\Big(\limsup_{t\to\infty} Z_\lambda(t) = \infty\Big) = 1.$$

□

Part III

Large-deviations via spines

Chapter 5

A large-deviations result for BBM

5.1 Introduction and outline of proof

Suppose that under a measure $\tilde{P}$ the process $(\xi_t)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}$. The large-deviations behaviour of ξ_t is controlled by Schilder's theorem, and in order to state this we first define a re-scaling of the paths down to the time-interval $s \in [0,1]$:

Definition 5.1.1 *If $(\xi_t)_{0\leq t\leq T}$ is the path in $\mathbb{R}$ followed over the time interval $t \in [0,T]$, then we define $(\xi^T(s))_{0\leq s\leq 1}$ to be a scaled-down version of this path:*

$$\xi^T(s) := T^{-1}\xi_{sT},$$

and refer to this ξ^T as the time-T re-scaled path.

Such a scaling of the path is effectively equivalent to supposing that ξ^T is a Brownian motion on $[0,1]$ with variance $1/\sqrt{T}$, and in fact Varadhan's proof of Schilder's theorem (see [55]) deals instead with a Brownian motion on $[0,1]$ whose diffusion coefficient is $\varepsilon > 0$ under a measure P_ε; thus he considers $\varepsilon \to 0$ rather than $T \to \infty$ and we could say heuristically that his ε is our $1/\sqrt{T}$. In our choice of the re-scaling approach we are following the large-deviations work of Git [15].

We use the label $C[0,1]$ to refer to the set of all continuous functions on $[0,1]$, and without losing generality we can suppose that under $\tilde{P}$ the Brownian motion starts at the origin.

Theorem 5.1.2 (Schilder) *There is a large-deviation principle for Brownian motion:*

- *Upper bound: If C is a closed subset of $C[0,1]$ then*

$$\limsup_{T\to\infty} T^{-1}\log\tilde{P}(\xi^T \in C) \leq -\inf_{g\in C} I(g),$$

- *Lower bound: If V is an open subset of $C[0,1]$ then*

$$\liminf_{T\to\infty} T^{-1}\log\tilde{P}(\xi^T\in V)\geq -\inf_{g\in V} I(g),$$

where

$$I(g) := \int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s$$

if $g\in C[0,1]$ with $g(0)=0$ has a square-integrable derivative; otherwise we define $I(g)=\infty$.

Now suppose that the $\mathbb{R}$-valued point process $\mathbb{X}_t := \{X_u(t) : u\in N_t\}_{t\geq 0}$ is a *branching* Brownian motion with constant branching rate r and binary-splitting, where N_t is the set of individuals alive at time t. We can likewise define a re-scaling of the paths:

Definition 5.1.3 *For each $T\geq 0$ and each $u\in N_T$ with path $X_u:[0,T]\to\mathbb{R}$, we define the function X_u^T on $[0,1]$ to be the* time-T re-scaled path*:*

$$X_u^T:[0,1]\to\mathbb{R},\quad X_u^T(s)=T^{-1}X_u(sT).$$

We remember that the particle u is born at the time $S_u-\sigma_u$, and for times earlier than this we interpret $X_u(t)$ naturally as the spatial position of the unique ancestor of u that was alive at time t – see Definition 3.1.2. As in previous chapters, we suppose that the probabilities of this BBM are $\{P^x : x\in\mathbb{R}\}$ so that P^x is a measure defined on the natural filtration $(\mathcal{F}_T)_{T\geq 0}$ such that it is the law of the process initiated from a single particle positioned at x. Without loss of generality we suppose that the initial ancestor of the BBM starts out at the origin, and henceforth use P to mean P^0.

Definition 5.1.4 *We use $C_0[0,1]$ to mean the set of paths $g\in C[0,1]$ with $g(0)=0$ whose derivative is square-integrable.*

In this chapter we are going to prove the following theorem concerning the probability that the path of *at least one* of the many particles in the branching diffusion stays near to a given continuous function.

Theorem 5.1.5 *There is a large-deviation result for the paths of a BBM:*

- *Upper bound: If C is a closed subset of $C[0,1]$ then*

$$\limsup_{T\to\infty} T^{-1}\log P(\exists u\in N_T : X_u^T\in C)\leq -\inf_{g\in C} S(g), \tag{5.1}$$

- *Lower bound: If V is an open subset of $C[0,1]$ then*

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u\in N_T : X_u^T\in V)\geq -\inf_{g\in V} S(g), \tag{5.2}$$

where

$$S(g) := \begin{cases} \sup_{w\in[0,1]}\left(\int_0^w \frac{1}{2}g'(s)^2 - r\;\mathrm{d}s\right) & \text{if } g\in C_0[0,1],\\ \infty & \text{otherwise.}\end{cases}$$

This large-deviations result was proven by Tzong-Yow Lee [37], but he relied heavily on Freidlin's previous work on rescalings of solutions of reaction-diffusion equations. Our proofs are based on spines, and offer much clearer, neater and independent proofs that can be generalized to cover many different types of branching diffusions – in the following chapter we develop the ideas to deal with the typed branching diffusion originally studied in Harris and Williams [21].

Strictly speaking the above result is not a *Large-Deviations Principle* in the precise sense – though we occasionally use this term in the following – since the probabilities $P(\exists u \in N_T : X_u^T \in \cdot)$ do not combine additively but are only *sub*-additive, and therefore cannot define a measure on $C[0,1]$. For example, if we choose a large $T > 0$ and define a set function $\mu_T : C[0,1] \to [0,1]$ via

$$\mu_T(A) := P(\exists u \in N_T : X_u^T \in A),$$

then it is clear that this is sub-additive:

$$\forall A, A_1, A_2 \in C[0,1] \text{ with } A \subset A_1 \cup A_2, \quad \mu_T(A) \leq \mu_T(A_1) + \mu_T(A_2).$$

To see that it is not additive, consider the two sets $A_1 := \{g \in C[0,1] : g(1) < 0\}$ and $A_2 := \{g \in C[0,1] : g(1) \geq 0\}$, for which $A_1 \cup A_2 = C[0,1]$. Then when T is large enough so that there are many Brownian particles in the BBM, we are very likely to find at least one particle has a path in A_1 and at least one other in A_2, to give:

$$\mu_T(A_1) + \mu_T(A_2) \simeq 1 + 1,$$

whilst $\mu_T(A_1 \cup A_2) = 1$.

The above comments are included just to make it clear that in this part of the thesis one must be careful when referring to so-called 'standard arguments' from large-deviations theory. To reassure the reader, in section 5.6 we are particularly careful to give full proofs of the topological properties that we use – but the reader can verify that the proofs we give closely mirror those given by Dembo and Zeitouni [12], which in fact depend only on this above property of sub-additivity at two points: corresponding to our equations (5.24) and (5.27).

We note that for some paths g we shall have $S(g) = 0$: for example if $g(s) = \lambda s$ with $\lambda^2 < 2r$. The large-deviations lower bound will then suggest that there is always a probability that a BBM path of this shape s present. In fact a much stronger result has been proven by Git [15] which essentially states that *almost surely* we can be sure to have not just one of these paths with $S(g) = 0$ present in the BBM but an *exponentially growing number.*

Outline of proof

As far as the topological issues in our arguments are concerned, the main reference is Dembo and Zeitouni [12]. It is known that the δ-neighbourhoods make up a base for the topology of $C[0,1]$ induced by the metric $\|f\| := \sup_{w \in [0,1]} |f(w)|$.

Definition 5.1.6 *For a given $g \in C[0,1]$ and $\delta > 0$ we define*

$$\mathbb{B}_\delta(g) := \{f \in C[0,1] : \|f - g\| < \delta\},$$

as the δ-neighbourhood around the function g.

We aim to prove Theorem 5.1.5 in two stages. First we use spine techniques to prove the following local results:

Theorem 5.1.7 *For any fixed $g \in C[0,1]$ we have*

$$\lim_{\delta \to 0} \limsup_{T \to \infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g); \tag{5.3}$$

$$\lim_{\delta \to 0} \liminf_{T \to \infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g). \tag{5.4}$$

For the local upper bound (5.3) we use the Many-to-One theorem to reduce the question to just the spine and then use Schilder's theorem. Our proof of the local lower bound (5.4) is based on getting an upper-bound for a new additive martingale for BBM, and we can illustrate the principle behind this with the BBM martingale $Z_\lambda(t)$ defined at (4.1): the fact that $\mathrm{d}\mathbb{Q}_\lambda/\mathrm{d}P = Z_\lambda(t)$ on $\mathcal{F}_t$ means that for a set $F \in \mathcal{F}_t$

$$P(F) = \mathbb{Q}_\lambda\Big(\frac{1}{Z_\lambda(t)}; F\Big),$$

and therefore an upper-bound on the growth of $Z_\lambda(t)$ under $\mathbb{Q}_\lambda$, (where it is a submartingale as we have seen) will here give us a lower bound for the probability $P(F)$. The main work of the local lower bound is therefore to define the correct martingale to give the appropriate new measure and to obtain a suitable upper-bound for it under this new measure – here the techniques we developed in chapter 4 are very useful, and we comment that without the new spine approach that we lay out here, such lower-bounds of large-deviation probabilities are notoriously difficult.

A topological-type theorem from Dembo and Zeitouni means that these two local results imply the existence of (at worst) a *weak* LDP – if we use this term also for the sub-additive probabilities of Theorem 5.1.5; this is just to say that the lower bound holds in full for any open set $V \subset C[0,1]$ but that the upper bound requires $C \subset C[0,1]$ to be closed and *compact.* Once Theorem 5.1.7 has been proved, in section 5.6 we use the Many-to-One theorem with the concept of *exponential tightness* for a single Brownian motion (the spine) to improve these local results to the full large-deviations statement of Theorem 5.1.5.

5.2 A local upper bound

Theorem 5.2.1 *Let $g \in C[0,1]$. Then,*

$$\lim_{\delta \to 0} \limsup_{T \to \infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \le -S(g). \tag{5.5}$$

Proof: We first note that a monotonocity holds:

$$0 \ge \limsup_{T \to \infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \downarrow \qquad \text{as } \delta \downarrow 0.$$

and therefore the $\delta \to 0$ limit (5.5) exists (though it could potentially be $-\infty$).

The probability that a single particle has a path near g is smaller than the *expected number* of such particles:

$$P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \le P\Big(\sum_{u\in N_t} \mathbf{1}\{X_u^T \in \mathbb{B}_\delta(g)\}\Big),$$

and an application of the Many-to-One theorem 3.6.5 gives:

$$\begin{aligned} P\Big(\sum_{u\in N_T} \mathbf{1}\{X_u^T \in \mathbb{B}_\delta(g)\}\Big) &= \tilde{P}\big(e^{rT}\mathbf{1}\{\xi^T \in \mathbb{B}_\delta(g)\}\big) \\ &= e^{rT} P\big(\xi^T \in \mathbb{B}_\delta(g)\big). \end{aligned} \tag{5.6}$$

If it is the case that $g \notin C_0[0,1]$, so that its derivative is not square-integrable, then from the simple fact that the open set $\mathbb{B}_\delta(g)$ is a subset of the closed δ-neighbourhood $\overline{\mathbb{B}_\delta(g)}$, we can use the above reasoning to deduce that

$$\begin{aligned} \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) &\le \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \overline{\mathbb{B}_\delta(g)}) \\ &\le e^{rT} P\big(\xi^T \in \overline{\mathbb{B}_\delta(g)}\big), \end{aligned}$$

and an application of the upper bound in Schilder's theorem to the right-hand probability will give us the correct result:

$$\lim_{\delta\to 0} \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \le r - \lim_{\delta\to 0} \inf_{g\in\overline{\mathbb{B}_\delta(g)}} I(g) = -\infty = -S(g).$$

Therefore we can assume that throughout the following proof we have the more interesting case of $g \in C_0[0,1]$.

The reasoning that gave (5.6) can immediately be strengthened by the simple observation that if the rescaled path is near g throughout the whole interval $[0,1]$, then it *must* be near g throughout all shorter intervals $[0,w]$, and a similar argument to the above would imply that for all $w \in [0,1]$,

$$\begin{aligned} P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) &\le P(\exists u \in N_T : \big|X_u^T(s) - g(s)\big| < \delta, \forall s \in [0,w]) \\ &\le P\Big(\sum_{u\in N_T} \mathbf{1}\{\big|X_u^T(s) - g(s)\big| < \delta, \forall s \in [0,w]\}\Big) \\ &= e^{rwT} \tilde{P}\big(\big|\xi^T(s) - g(s)\big| < \delta, \forall s \in [0,w]\big). \end{aligned} \tag{5.7}$$

For $g \in C_0[0,1]$ it is clear that the supremum in the definition of the rate functional $S(g)$ will be reached at some point $\hat{w} \in [0,1]$, whence

$$S(g) = -r\hat{w} + \int_0^{\hat{w}} \frac{1}{2} g'(s)^2 \, \mathrm{d}s.$$

Choosing $w = \hat{w}$ in (5.7) gives

$$\begin{aligned} &\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \\ &\qquad \le r\hat{w} + \limsup_{T\to\infty} T^{-1} \log \tilde{P}\big(\big|\xi^T(s) - g(s)\big| < \delta, \forall s \in [0,\hat{w}]\big). \end{aligned} \tag{5.8}$$

Since the spine diffusion ξ_t is just a Brownian motion, Schilder's theorem says that over the time interval $[0, \hat{w}]$, its re-scaled path $\xi^T(s)$ will satisfy a large-deviations principle with rate functional $I^{\hat{w}}(g) := \frac{1}{2}\int_0^{\hat{w}} g'(s)^2 \, \mathrm{d}s$, and therefore,

$$\lim_{\delta\to 0} \limsup_{T\to\infty} T^{-1} \log \tilde{P}\big(\big|\xi^T(s) - g(s)\big| < \delta, \forall s \in [0, \hat{w}]\big) = -\frac{1}{2}\int_0^{\hat{w}} g'(s)^2 \, \mathrm{d}s. \tag{5.9}$$

Our local upper bound for the BBM now follows directly from (5.8) and (5.9):

$$\begin{aligned}\lim_{\delta\to 0} \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) &\leq r\hat{w} - \frac{1}{2}\int_0^{\hat{w}} g'(s)^2 \, \mathrm{d}s \\ &= -S(g).\end{aligned}$$

□

5.3 A new martingale Z_{g_T} for BBM

Let $g \in C_0[0,1]$ be a fixed path; note that from here until Section 5.5 we must insist that the derivative g' is square-integrable, since otherwise the change-of-measure martingales cannot be defined.

Given that the spine diffusion ξ_t is itself a $\tilde{P}$-Brownian motion, it follows that on the sub-filtration $(\mathcal{G}_t)_{0\leq t\leq T}$,

$$\zeta_{g_T}(t) := \exp\Big\{\int_0^t g_T'(s) \, \mathrm{d}\xi_s - \frac{1}{2}\int_0^t g_T'(s)^2 \, \mathrm{d}s\Big\}, \tag{5.10}$$

is a $\tilde{P}$-martingale, where

Definition 5.3.1 *for any fixed $T \geq 0$ and any function $g \in C[0,1]$ we define*

$$g_T(s) := Tg(s/T) \qquad \forall s \in [0, T].$$

to be the time-T scaled up version of g.

This martingale (5.10) is well known from the Girsanov theorem, and when used to change the measure it will introduce a drift to the Brownian motion.

Likewise, the process n_t which counts the number of fission times on the spine up to time t is a Poisson process of rate r and therefore

$$t \mapsto e^{-rt} 2^{n_t}$$

is a $\tilde{P}$-martingale too which will increase the rate of n_t from r to $2r$ if used to change the measure – see Kyprianou [35].

We use the product of these two martingales to define a new measure:

Theorem 5.3.2 *For each $T \geq 0$ we define a measure $\tilde{\mathbb{Q}}_T$ on the filtration $(\tilde{\mathcal{F}}_t)_{0\leq t\leq T}$ via*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_T}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_t} = e^{-rt}2^{n_t} \times \zeta_{g_T}(t). \tag{5.11}$$

Under the measure $\tilde{\mathbb{Q}}_T$ we can give a pathwise construction of the branching-diffusion $\mathbb{X}_t$ over the time-interval $t \in [0, T]$:

- *the spine diffusion $(\xi_t)_{0\leq t\leq T}$ starts at 0 and diffuses so $\xi_t - g_T(t)$ is a $\tilde{\mathbb{Q}}_T$-Brownian motion over the time interval $t \in [0, T]$;*
- *at rate $2r$ the spine undergoes fission producing two particles;*
- *with equal probability, one of these two particles is selected to continue the spine;*
- *the other particle initiates, from its birth position, an independent copy of a $P^{\cdot}$ branching Brownian motion with branching rate r.*

We briefly recall that for the third point above, the two particles produced are born at the same location and are therefore spatially indistinguishable, but due to our use of the Ulam-Harris labelling scheme they are distinguishable according to the label that they carry. Therefore the idea of choosing a particle is really a question of choosing between *labels.*

As we have seen before, this change of measure gives us an additive martingale over the branching particles:

Definition 5.3.3 *For each $T \geq 0$,*

$$Z_{g_T}(t) := e^{-rt} \sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s) - \frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s},$$

defines an additive martingale on the filtered probability space $\left(\tilde{T}, \mathcal{F}_\infty, (\mathcal{F}_t)_{0\leq t\leq T}\right)$.

That this is really a martingale is due to the following:

Theorem 5.3.4 *If we define $\mathbb{Q}_{g_T} := \tilde{\mathbb{Q}}_{g_T}|_{\mathcal{F}_T}$, then $\mathbb{Q}_{g_T}$ is a measure on the filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ and*

$$\left.\frac{\mathrm{d}\mathbb{Q}_T}{\mathrm{d}P}\right|_{\mathcal{F}_t} = Z_{g_T}(t).$$

Proof: It is clear from the definition of the conditional expectation that the change of measure (5.11) projects onto the sub-algebra $\mathcal{F}_t$ as a conditional expectation: for $t \in [0, T]$

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_T}{\mathrm{d}\tilde{P}}\right|_{\mathcal{F}_t} = \tilde{P}\left(e^{-rt}2^{n_t}\zeta_{g_T}(t)|\mathcal{F}_t\right).$$

Bearing in mind that $2^{n_t} = \prod_{v<\xi_t} 2$, if we use the representation (3.4) we get

$$\begin{aligned}
\tilde{P}\big(e^{-rt}2^{n_t}\zeta_{g_T}(t)|\mathcal{F}_t\big) &= \tilde{P}\Big(e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2 \times \mathbf{1}_{(\xi_t=u)}|\mathcal{F}_t\Big) \\
&= e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2 \times \tilde{P}(\xi_t = u|\mathcal{F}_t) \\
&= e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} \times \prod_{v<u} 2 \times \prod_{v<u}\frac{1}{2} \\
&= e^{-rt}\sum_{u\in N_t} e^{\int_0^t g_T'(s)\,\mathrm{d}X_u(s)-\frac{1}{2}\int_0^t g_T'(s)^2\,\mathrm{d}s} = Z_{g_T}(t).
\end{aligned}$$

We note that in this above chain we used $\tilde{P}\big(\xi_t = u|\mathcal{F}_t\big) = \prod_{v<u}\frac{1}{2}$, which is the case of (3.3) for binary-splitting. □

5.4 The growth of Z_{g_T} under $\mathbb{Q}_{g_T}$

We are going to use the *spine decomposition* to get a good estimate of $\mathbb{Q}_T(Z_{g_T}(T)^\alpha)$ that we can use in Doob's submartingale inequality.

Theorem 5.4.1 *For each $g \in C_0[0,1]$ and for each $\alpha \in [0,1]$,*

$$\mathbb{Q}_T\big(Z_{g_T}(T)^\alpha\big) \leq e^{\alpha S(g)T} e^{\frac{1}{2}\alpha^2 T\int_0^1 g'(s)^2\,\mathrm{d}s}\big(1+2rT\big). \tag{5.12}$$

Proof: Since it is only the spine that is affected by the change of measure, the so-called *spine decomposition* in which we condition on knowing the spine's behaviour and fission times, is exceptionally useful for dealing with the P-martingale:

$$\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)|\tilde{\mathcal{G}}_\infty\big) = e^{-rT}e^{\int_0^T g_T'(s)\,\mathrm{d}\xi_s-\frac{1}{2}\int_0^T[g_T'(s)]^2\,\mathrm{d}s} + \sum_{u<\xi_T} e^{-rS_u}e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\xi_s-\frac{1}{2}\int_0^{S_u}[g_T'(s)]^2\,\mathrm{d}s}. \tag{5.13}$$

We recall that the filtration $\tilde{\mathcal{G}}_\infty$ contains all information about the spine and the fission times S_u that occur along it.

By definition, $\hat{\xi}_s := \xi_s - g_T(s)$ for $0 \leq s \leq T$ is a Brownian motion under $\tilde{\mathbb{Q}}_T$, and substituting

$$\mathrm{d}\xi_s = \mathrm{d}\hat{\xi}_s + g_T'(s)\mathrm{d}s, \tag{5.14}$$

into (5.13) we arrive at:

$$\begin{aligned}
\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)|\tilde{\mathcal{G}}_\infty\big) &= e^{\frac{1}{2}\int_0^T g_T'(s)^2\,\mathrm{d}s-rT}e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\frac{1}{2}\int_0^{S_u} g_T'(s)^2\,\mathrm{d}s-rS_u}e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s} \\
&= e^{\left(\int_0^1 \frac{1}{2}g'(s)^2-r\,\mathrm{d}s\right)T}e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\left(\frac{1}{2}\int_0^{S_u/T}[g'(s)]^2-r\,\mathrm{d}s\right)S_u}e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s} \\
&\leq e^{\left(\sup_{w\in[0,1]}\int_0^w \frac{1}{2}g'(s)^2-r\,\mathrm{d}s\right)T}\left(e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\right), \\
&= e^{S(g)T}\left(e^{\int_0^T g_T'(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\int_0^{S_u} g_T'(s)\,\mathrm{d}\hat{\xi}_s}\right).
\end{aligned} \tag{5.15}$$

In the above we note that $g'_T(s) = g'(s/T)$.

From the tower property, and since $\mathbb{Q}_T = \tilde{\mathbb{Q}}_T$ on $\mathcal{F}_T$,

$$\mathbb{Q}_T\big(Z_{g_T}(T)^\alpha\big) = \tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)^\alpha\big) = \tilde{\mathbb{Q}}_T\Big(\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)^\alpha|\tilde{\mathcal{G}}_\infty\big)\Big),$$

and the conditional form of Jensen's inequality says that for $\alpha \in [0,1]$,

$$\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)^\alpha|\tilde{\mathcal{G}}_\infty\big) \le \tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)|\tilde{\mathcal{G}}_\infty\big)^\alpha.$$

Since the spine decomposition $\tilde{\mathbb{Q}}_T\big(Z_{g_T}(T)|\tilde{\mathcal{G}}_\infty\big)$ is a sum, we can use the following result noted by Neveu [44]

Proposition 5.4.2 *If $\alpha \in (0,1]$ and $u, v > 0$ then $(u+v)^\alpha \le u^\alpha + v^\alpha$.*

Continuing from (5.15) these lead to

$$\mathbb{Q}_T\big(Z_{g_T}(T)^\alpha\big) \le e^{\alpha S(g)T}\,\tilde{\mathbb{Q}}_T\Big(e^{\alpha\int_0^T g'_T(s)\,\mathrm{d}\hat{\xi}_s} + \sum_{u<\xi_T} e^{\alpha\int_0^{S_u} g'_T(s)\,\mathrm{d}\hat{\xi}_s}\Big). \tag{5.16}$$

Under the measure $\tilde{\mathbb{Q}}_T$, the process $(\hat{\xi}_t)_{0\le t\le T}$ is a standard Brownian motion, and therefore

$$e^{\alpha\int_0^T g'_T(s)\,\mathrm{d}\hat{\xi}_s - \frac{1}{2}\alpha^2\int_0^T g'_T(s)^2\,\mathrm{d}s}$$

is a $\tilde{\mathbb{Q}}_T$-martingale on $t \in [0,T]$. Evaluating this at the bounded stopping-times $\big(S_u : u < \xi_T\big)$ gives

$$\tilde{\mathbb{Q}}_T\big(e^{\alpha\int_0^{S_u} g'_T(s)\,\mathrm{d}\hat{\xi}_s}\big) = \tilde{\mathbb{Q}}_T\big(e^{\frac{1}{2}\alpha^2\int_0^{S_u} g'_T(s)^2\,\mathrm{d}s}\big) \le e^{\frac{1}{2}\alpha^2\int_0^T g'_T(s)^2\,\mathrm{d}s},$$

whence from (5.16) we obtain

$$\mathbb{Q}_T\big(Z_{g_T}(T)^\alpha\big) \le e^{\alpha S(g)T} e^{\frac{1}{2}\alpha^2\int_0^T g'_T(s)^2\,\mathrm{d}s}\,\tilde{\mathbb{Q}}_T\big(1+n_T\big).$$

We know that under the measure $\tilde{\mathbb{Q}}_T$ the births on the spine occur as a Poisson process with rate $2r$, whence the expectation grows linearly in T:

$$\tilde{\mathbb{Q}}_T\big(1+n_T\big) = 1 + 2rT,$$

and we arrive at

$$\mathbb{Q}_T\big(Z_{g_T}(T)^\alpha\big) \le e^{\alpha S(g)T} e^{\frac{1}{2}\alpha^2 T\int_0^1 g'(s)^2\,\mathrm{d}s}\big(1+2rT\big).$$

□

Theorem 5.4.3 *For each $\varepsilon > 0$,*

$$\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) \le e^{(S(g)+\varepsilon)T}\Big) \to 1, \qquad \textit{as } T \to \infty. \tag{5.17}$$

Proof: First of all, for each $\alpha \in [0,1]$, a proof very similar to that given in Theorem 4.2.9 implies that $Z_{g_T}(t)^\alpha$ is a $\mathbb{Q}_T$ submartingale on $t \in [0,T]$, and we now prove a probability bound on its growth.

The estimate (5.12) works well with Doob's submartingale inequality: for any small $\varepsilon > 0$ and for any fixed $T > 0$,

$$\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_\lambda(s) > e^{(S(g)+\varepsilon)T}\Big) = \mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_\lambda(s)^\alpha > e^{\alpha(S(g)+\varepsilon)T}\Big) \le \frac{\mathbb{Q}_T(Z_\lambda(T)^\alpha)}{e^{\alpha(S(g)+\varepsilon)T}}.$$

Using (5.12) this gives

$$\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) > e^{(S(g)+\varepsilon)T}\Big) \le e^{\left(\alpha\int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s - \varepsilon\right)\alpha T}\big(1+2rT\big).$$

Bearing in mind that $\int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s$ is just a finite number, we can choose $\alpha > 0$ small enough so that $\alpha\int_0^1 \frac{1}{2}g'(s)^2\,\mathrm{d}s - \varepsilon < 0$, whence the exponential decay dominates the linear growth in the above, and we have proven that

$$\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) > e^{(S(g)+\varepsilon)T}\Big) \to 0, \qquad \text{as } T \to \infty.$$

$\square$

5.5 A local lower bound

We note that in the case of the lim inf we do not deal immediately with the limit as $\delta \to 0$, since without the monotonicity that we had for the lim sup we do not *a priori* know that the limit exists.

Theorem 5.5.1 *Let $g \in C[0,1]$. For any fixed $\delta > 0$, we have*

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) \ge -S(g). \tag{5.18}$$

Proof: First we note that if g is not in $C_0[0,1]$ then $S(g) = \infty$ and the result holds trivially. Therefore we assume that throughout we have $g \in C_0[0,1]$.

Importantly, the event we are considering is $\mathcal{F}_T$-measurable, and on this σ-algebra the change of measure is carried out by Z_{g_T}, as stated in Theorem 5.3.4. Therefore,

$$P\big(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)\big) = \mathbb{Q}_T\Big(\frac{1}{Z_{g_T}(T)}; \exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)\Big). \tag{5.19}$$

The upper bound that we have derived for Z_{g_T} will now serve as a lower bound for $1/Z_{g_T}(T)$, so that for any $\varepsilon > 0$,

$$\begin{aligned} &P\big(\exists u \in N_T\colon X_u^T \in \mathbb{B}_\delta(g)\big) \\ &\ge e^{-(S(g)+\varepsilon)T}\mathbb{Q}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) \le e^{(S(g)+\varepsilon)T}; \exists u \in N_T\colon X_u^T \in \mathbb{B}_\delta(g)\Big) \\ &\ge e^{-(S(g)+\varepsilon)T}\tilde{\mathbb{Q}}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) \le e^{(S(g)+\varepsilon)T};\ \xi^T \in \mathbb{B}_\delta(g)\Big). \end{aligned} \tag{5.20}$$

Since $\xi^T_s - g(s)$ is a $\tilde{\mathbb{Q}}_T$-Brownian motion on $[0,1]$ with diffusion coefficient $1/\sqrt{T}$, it follows that

$$\tilde{\mathbb{Q}}_T(\xi^T \in \mathbb{B}_\delta(g)) \to 1, \qquad \text{as } T \to \infty,$$

and this combines with the result of Theorem 5.4.3 to give:

$$\tilde{\mathbb{Q}}_T\Big(\sup_{s\in[0,T]} Z_{g_T}(s) \le e^{(S(g)+\varepsilon)T};\ \xi^T \in \mathbb{B}_\delta(g)\Big) \to 1, \qquad \text{as } T \to \infty.$$

Thus from (5.20) we have

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X^T_u \in \mathbb{B}_\delta(g)) \ge -S(g) - \varepsilon$$

which proves (5.18) since ε was arbitrary. □

Corollary 5.5.2 *For each $g \in C[0,1]$ we have*

$$\lim_{\delta\to 0} \liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X^T_u \in \mathbb{B}_\delta(g)) = -S(g). \tag{5.21}$$

Proof: The case where $g \notin C_0[0,1]$ is immediate from the above, and therefore we assume that $g \in C_0[0,1]$. We have proved Theorem 5.2.1 which can be interpreted as saying that for each $\delta > 0$ there is an $\varepsilon_\delta > 0$ such that

$$-S(g) + \varepsilon_\delta > \limsup_{t\to\infty} T^{-1} \log P(\exists u \in N_T : X^T_u \in \mathbb{B}_\delta(g)),$$

with $\varepsilon_\delta \to 0$ as $\delta \to 0$. Now a trivial inequality between the limsup and liminf combines with this and with the lower bound of Theorem 5.18 to give

$$\begin{aligned}-S(g) + \varepsilon_\delta &> \limsup_{t\to\infty} T^{-1} \log P(\exists u \in N_T : X^T_u \in \mathbb{B}_\delta(g)) \\ &\ge \liminf_{t\to\infty} T^{-1} \log P(A^{g,\delta}_T) \ge -S(g),\end{aligned}$$

which implies that the $\delta \to 0$ limit exists also for the lim inf. □

Together with Theorem 5.2.1 we have now completed the proof of the local limit result Theorem 5.1.7.

5.6 Improving the 'weak' large-deviations result

As mentioned, the local results of Theorem 5.1.7 and the fact that the δ-neighbourhoods $\mathbb{B}_\delta(g)$ form a base for the topology of $C[0,1]$ means that we have at least a *weak* large-deviations result: the lower bound of Theorem 5.1.5 holds, but the upper bound is proven only for *compact* sets (as opposed to *closed* sets). The main ideas for the following proof of this come from Theorem 4.1.11 of Dembo and Zeitouni [12].

Theorem 5.6.1 *The local results of Theorem 5.1.7 imply that the upper bound of our main result Theorem 5.1.5 holds certainly for all $C \subset C[0,1]$ that are closed and* compact*:*

$$\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in C) \le -\inf_{g\in C} S(g),$$

whilst the lower bound holds in full for all open subsets $V \subset C[0,1]$:

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in V) \ge -\inf_{g\in V} S(g).$$

Proof: First of all we consider the lower bound. The open δ-neighbourhoods $\{\mathbb{B}_\delta(g) : g \in C[0,1], \delta > 0\}$ form a base for the topology of $C[0,1]$ which we shall call $\mathcal{A}$. Therefore if $V \subset C[0,1]$ is an open set then for each $g \in V$ we can be sure that for some small enough $\delta > 0$ we shall have $g \in \mathbb{B}_\delta(g) \subset V$, and therefore

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in V) \ge \liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)).$$

Furthermore, since the δ-neighbourhoods sit inside one another as $\delta \to 0$ we actually have a limit result which combines with result (5.21) to say that for each $g \in V$:

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in V) \ge \lim_{\delta\to 0}\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g).$$

Since this holds for all $g \in V$ it will hold for the supremum:

$$\liminf_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in V) \ge \sup_{g\in V} -S(g) = -\inf_{g\in V} S(g),$$

concluding the proof of the lower bound.

For the upper bound we use a finite-covering argument: supposing that $C \subset C[0,1]$ is closed and compact we shall cover it with a finite number of open sets from $\mathcal{A}$ to deduce the result. Theorem 5.5 states that

$$\lim_{\delta\to 0}\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in \mathbb{B}_\delta(g)) = -S(g).$$

Since each open set $A \in \mathcal{A}$ contains at least one δ-neighbourhood, this result implies that

$$\inf_{\{A\in\mathcal{A}, g\in A\}} \limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in A) = -S(g),$$

and for the following argument we re-arrange this as

$$\sup_{\{A\in\mathcal{A}, g\in A\}} \Big[-\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in A)\Big] = S(g). \tag{5.22}$$

If we choose and fix $\delta > 0$ and define

$$S^\delta(g) := \min\{S(g) - \delta, 1/\delta\},$$

then for each $g \in C[0,1]$ the above (5.22) implies that there is some open set $A_g \in \mathcal{A}$ (which may depend on δ) such that

$$-\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in A_g) \ge S^\delta(g). \tag{5.23}$$

For the given compact set C we can extract a finite cover from the covering $\bigcup_{g\in C} A_g$, which we denote $\{A_{g_1}, \dots A_{g_n}\}$; then by the sub-additivity property discussed earlier in this chapter we have

$$P(\exists u \in N_T : X_u^T \in C) \le \sum_{i=1}^{n} P(\exists u \in N_T : X_u^T \in A_{g_i}). \tag{5.24}$$

Here we are dealing with a finite sum and can use a standard result of Laplace on the growth rate of finite sums of exponentials (see Dembo and Zeitouni's Lemma 1.2.15):

$$\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in C) \le \max_{i=1,\dots,n} \limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in A_{g_i}).$$

From (5.23) we have

$$\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in A_{g_i}) \le -S^\delta(g_i)$$

and therefore

$$\limsup_{T\to\infty} T^{-1}\log P(\exists u \in N_T : X_u^T \in C) \le -\min_{i=1,\dots,n} S^\delta(g_i) \le -\inf_{g\in C} S^\delta(g).$$

The proof of the upper bound is completed by considering the limit as $\delta \to 0$. □

It is clear how the compactness property was the connecting link between the local properties of Theorem 5.1.7 and the above *weak* large-deviations result. We now wish to improve this to get the full large-deviations result of Theorem 5.1.5, and a neat way to do this is to use a property of measures known as *exponential tightness.* This approach is particularly suitable for spines since the question of exponential tightness of the BBM probabilities can be reduced to that of the plain Brownian-motion probabilities, by the Many-to-One theorem.

Definition 5.6.2 *A family of probability measures $\{\mu_T\}$ on a set $\mathcal{X}$ is said to be exponentially tight if for each $\alpha < \infty$ there exists a compact $K \subset \mathcal{X}$ such that*

$$\limsup_{T\to\infty} T^{-1}\log \mu_T(K^{\mathsf{C}}) < -\alpha,$$

where K^{C} denotes the set complement.

We recall without proof a standard result that on the set of paths $C[0,1]$, the measures

$$P(\xi_T \in A), \qquad \text{for } A \subset C[0,1],$$

concerning the paths of a single Brownian motion ξ_T are exponentially tight; a remark to this effect can be found on page 120 of Dembo and Zeitouni.

Theorem 5.6.3 *The fact that the above path-measures of a single Brownian motion are exponentially tight implies that the (sub-additive) measures $\{P(\exists u \in N_T : X_u^T \in \cdot\,)\}_{T\ge 0}$ for the branching Brownian motion are also exponentially tight.*

Proof: For any set $K \subset C[0,1]$, we have an expectation bound that combines with the Many-to-One property to give:

$$P(\exists u \in N_T : X_u^T \in K^{\complement}) \le P\Big(\sum_{u \in N_T} \mathbf{1}\{X_u^T \in K^{\complement}\}\Big) = e^{rT} P(\xi^T \in K^{\complement}),$$

whence

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) \le r + \limsup_{T\to\infty} T^{-1} \log P(\xi^T \in K^{\complement}). \tag{5.25}$$

Let $\alpha < \infty$ be given. Since the spine is a Brownian motion, for which it is known that the probabilities $P(\xi^T \in \cdot\,)$ are exponentially tight, we can find some compact $K \subset C[0,1]$ such that

$$\limsup_{T\to\infty} T^{-1} \log P(\xi^T \in K^{\complement}) < -r - \alpha,$$

and therefore from (5.25),

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) < -\alpha.$$

□

Dembo and Zeitouni [12] state at Lemma 1.2.18, that when an exponentially tight family of measures satisfy a weak LDP then in fact the LDP holds in full. For completeness we state their proof as it applies to our particular context:

Theorem 5.6.4 *The weak LDP result of Theorem 5.6.1 together with the exponential tightness proven in Theorem 5.6.3 imply that the large-deviations result Theorem 5.1.5 of this chapter holds in full.*

Proof: The lower bound of Theorem 5.1.5 is exactly the same as that proven in the weak version of Theorem 5.6.1 and therefore here we are only looking to extend the upper bound of Theorem 5.6.1 for closed and compact sets to hold for the large class of all closed sets. That is, we want to show that for each closed subset $C \subset C[0,1]$ we have

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C) \le -\inf_{g \in C} S(g).$$

From the fact that we have exponential tightness of the probabilities, we know that there is some compact subset $K \subset C[0,1]$ such that

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement}) < -\inf_{g \in C} S(g); \tag{5.26}$$

we point out that we have just taken $\alpha = \inf_{g \in C} S(g)$ in Definition 5.6.2. The covering of C as

$$C = (C \cap K) \cup (C \cap K^{\complement}) \subset (C \cap K) \cup K^{\complement},$$

together with the sub-additivity property of our probabilities gives

$$P(\exists u \in N_T : X_u^T \in C) \le P(\exists u \in N_T : X_u^T \in C \cap K) + P(\exists u \in N_T : X_u^T \in K^{\complement}). \tag{5.27}$$

Now we are in a position to apply the 'weak' upper bound of Theorem 5.6.1 to the *compact* set $C \cap K$ to obtain

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C \cap K) \le -\inf_{g\in C\cap K} S(g) \le -\inf_{g\in C} S(g).$$

Applying the simple Laplace bound (Dembo and Zeitouni Lemma 1.2.15 as mentioned previously) to this above and (5.26) we obtain the desired result:

$$\begin{aligned}\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C) &\le \max\left\{\begin{array}{l}\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in C \cap K) \\ \limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\complement})\end{array}\right\} \\ &\le -\inf_{g\in C} S(g).\end{aligned}$$

□

This concludes the proof of the main result of this chapter, Theorem 5.1.5. We however include a final brief section on the rate function $S(g)$ for BBM.

5.7 A brief discussion of $S(g)$

$S(g)$ has been defined earlier as

$$S(g) := \begin{cases} \sup_{w\in[0,1]}\left(\int_0^w \frac{1}{2}g'(s)^2 - r \, \mathrm{d}s\right) & \text{if } g \in C_0[0,1], \\ \infty & \text{otherwise.}\end{cases}$$

In this section we give short proofs that this $S(g)$ is actually a so-called *good rate function.* Such facts were proven in Lee's [37] by arguments involving more heavy analytic estimates.

Theorem 5.7.1 *$S(g)$ is a good rate function. This is to say two things:*

- *it is a rate function, which is defined as stating that it is non-negative and that its level sets $\{g \in C[0,1] : S(g) \le \alpha\}$ are closed subsets of $C[0,1]$ for each α;*
- *it is a good rate function, which means that its level sets are actually compact subsets of $C[0,1]$.*

The ideas used in the following come from Dembo and Zeitouni [12].
Proof: First of all we trivially have $S(g) \ge 0$ for all $g \in C[0,1]$. Let $\alpha \ge 0$ be given. If $S(g) > \alpha$ then from (5.22) which says

$$S(g) = \sup_{\{A\in\mathcal{A}, g\in A\}} \left[-\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A)\right]$$

it follows that there is an open neighbourhood $A \in \mathcal{A}$ of g for which

$$-\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A) > \alpha.$$

This would therefore imply from (5.22) that for each $f \in A$ we also have

$$S(f) \geq -\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in A) > \alpha,$$

which is to say that $S(f) > \alpha$ for all $f \in A$. Hence the set $\{g : S(g) > \alpha\}$ is open and S really is a rate function.

To prove that S is a good rate function we use the lower bound of our large-deviations result together with the property of exponential tightness which says that for any given α there is some compact subset $K \subset C[0,1]$ for which

$$\limsup_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\mathsf{C}}) < -\alpha.$$

At the same time, since K^{C} is an open set we can apply the lower bound to get

$$\liminf_{T\to\infty} T^{-1} \log P(\exists u \in N_T : X_u^T \in K^{\mathsf{C}}) \geq -\inf_{g\in K^{\mathsf{C}}} S(g).$$

These two together force

$$\inf_{g\in K^{\mathsf{C}}} S(g) > \alpha,$$

from which we deduce that $\{g : S(g) \leq \alpha\} \subset K$. The level set is therefore a closed subset of a compact set K, whence it too is compact, and whence S is a good rate function. $\square$

Chapter 6

Large deviations for the HW model

6.1 Introduction and statement of result

We are going to use our spine techniques to prove an important large-deviations lower bound for the typed branching diffusion that we have already considered in section 4.4. We briefly remember that in this model, under a measure P, the type $Y_u(s)$ of a particle moves on the real line as an Orstein-Uhlenbeck process $\mathrm{OU}(\theta, \frac{\theta}{2})$ associated with the generator

$$Q_\theta := \frac{\theta}{2}\Big(\frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y}\Big), \qquad \text{for some fixed } \theta > 0,$$

whilst the spatial motion $X_u(s)$ of a particle of type $Y_u(s) = y$ is a driftless Brownian motion with instantaneous variance

$$A(y) := ay^2, \qquad \text{for some fixed } a > 0,$$

and fission of a particle of type y occurs at a rate

$$R(y) := ry^2 + \rho, \qquad \text{where } r, \rho > 0 \text{ are fixed},$$

to produce two particles at the same type-space location as the parent.

The precise form of the following result is motivated by the work carried out in Harris and Git [25], which we discuss in section 6.2. Actually, the spine techniques we use in proving this theorem do naturally give a much stronger result where particles not only arrive at the space-type location $(\beta t, \kappa\sqrt{t})$ at time τ, but are known to have stayed near a specific space-type path throughout the whole time interval $[0, \tau]$. We state this stronger result as Theorem 6.3.2, and here give the weaker form as it is required by the work of Harris and Git [25]:

Theorem 6.1.1 *Let $\tau > 0$ be fixed, and let $\beta, \kappa \in \mathbb{R}$ with $\beta < 0$ be given and fixed too. Then, for large t, the probability that at least one of the branching particles will be near $(\beta t, \kappa\sqrt{t})$*

at time τ (given that the original ancestor was at the space-type origin) has a large-deviations lower bound: for all $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \big|X_u(\tau) - \beta t\big| < \delta t, \big|Y_u(\tau) - \kappa\sqrt{t}\big| < \delta'\sqrt{t}\Big) \geq -\Theta(\beta, \kappa),$$

where

$$\Theta(\beta,\kappa) := \frac{\kappa^2}{4} + \frac{\sqrt{\theta(\theta - 8r)(a^2\kappa^4 + 4\theta\beta^2)}}{4a\theta}.$$

We note that as explained by Harris and Git, and as our spine proof will make clear, this lower-bound of $-\Theta(\beta, \kappa)$ is not exactly optimal and can be marginally improved to a rate of $-J(\tau)$, to be defined shortly; but anyway we would have $-J(\tau) \downarrow -\Theta(\beta, \kappa)$ as $\tau \to \infty$, and therefore for the work of Harris and Git there is not a real loss in the lower bound of $-\Theta(\beta, \kappa)$. As far as an upper bound is concerned, as is typically the case in large deviations it is generally easier to obtain than the lower bound. Here we could use the Many-to-One theorem to carry this out much like we did for the case of BBM in the previous chapter.

The principle behind the proof of the lower bound is to design new measures $\mathbb{Q}_t$ for the branching diffusion such that under $\mathbb{Q}_t$ one of the particles (the spine) will follow a specific space-type path to arrive at the point $(\beta t, \kappa\sqrt{t})$. Our theory laid out in chapter 3 will allow us to explicitly find the Radon-Nikodym derivatives (martingales) of these new measures with respect to the original measure P, and using the spine decomposition together with Doob's submartingale inequality we shall show that the growth rate of these martingales under $\mathbb{Q}_t$ is exactly the correct rate for the large-deviations.

In the next section we discuss the results of Harris and Git [25] in order to give a context to our work. The following section 6.3 contains a heuristic discussion of the large deviations for the model, and can be considered as the motivating arguments behind the choice of a martingale that will be used to carry out a change of measure. In section 6.4 these strictly-positive martingales Z_t are defined in terms of specific paths that our heuristics will have suggested, and we use the spine decomposition to get a good upper-bound on their growth in the new measures $\mathbb{Q}_t$ that they can define as Radon-Nikodym derivative (they become *sub*martingales for the new measure). This upper bound on Z_t under $\mathbb{Q}_t$ then gives us our lower bound for the P-probability of Theorem 6.1.1, in the final section 6.5.

6.2 Review of the Harris and Git almost-sure result

We summarize the main results from Harris and Git [25] to show how Theorem 6.1.1 fits into the picture. Work on almost-sure large-deviations results for this typed branching-diffusion actually began in the paper by Harris and Williams [21], and are continued in Harris and Git [25] where they proved that if we define a counting function

$$N_t(\gamma) = \sum_{u\in N_t} \mathbf{1}_{(X_u(t)\leq -\gamma t)}$$

for each $\gamma \in \mathbb{R}$, then the limit

$$\lim_{t\to\infty} t^{-1} \log N_t(\gamma) = \Delta(\gamma)$$

exists almost surely and is *finite* for all $0 \le \gamma < \tilde{c}(\theta)$, for the constant $\tilde{c}(\theta)$ defined in section 4.4 at lemma 4.4.1; in the case that $\gamma \ge \tilde{c}(\theta)$ the limit is $-\infty$ since *no* particles will be near such a spatial ray at large times. In other words, this result says that we almost-surely have exponential growth in numbers of particles following close to spatial rays that are *not too steep*. For later reference we state that

$$\Delta(\gamma) = \inf_{\lambda\in(\lambda_{\min},0)} \{E_\lambda^- + \lambda\gamma\} = \rho + \frac{\theta}{4} - \frac{1}{4}\sqrt{\theta(\theta-8r)(1+4\gamma^2/(\theta a))},$$

where $E_\lambda^- \in \mathbb{R}$ is the eigenvalue defined at equation 4.24, and

$$\tilde{c}(\theta) = \sup\{\gamma : \Delta(\gamma) > 0\} = \sqrt{2a\Big(r+\rho+\frac{2(2r+\rho)^2}{\theta-8r}\Big)}.$$

The work of Harris and Git [25] aims at improving this to obtain the almost-sure rate of growth in numbers of particles at certain *spatial and type* positions at large times. They define the following function that counts how many particles occupy a particular region in the type-space domain:

$$N_t(\gamma,\kappa) := \sum_{u\in N(t)} \mathbf{1}\{X_u(t) \le -\gamma t, Y_u(t)^2 \ge \kappa^2 t\}.$$

Harris and Git's work is directed at proving the following *almost-sure* result:

Theorem 6.2.1 *Under each $P^{x,y}$ law, the limit*

$$D(\gamma,\kappa) := \lim_{t\to\infty} t^{-1}\log N_t(\gamma,\kappa)$$

exists almost-surely and is given by

$$D(\gamma,\kappa) = \begin{cases} \Delta(\gamma,\kappa) & \text{if } \Delta(\gamma,\kappa) > 0, \\ -\infty & \text{otherwise.} \end{cases}$$

Here,

$$\begin{aligned} \Delta(\gamma,\kappa) &= \inf_{\lambda\in(\lambda_{min},0)} \{E_\lambda^- + \lambda\gamma - \kappa^2\psi_\lambda^+\}, \\ &= \rho + \frac{\theta-\kappa^2}{4} - \frac{1}{4a\theta}\sqrt{\theta(\theta-8r)(4a\theta\gamma^2 + a^2(\theta+\kappa^2)^2)}. \end{aligned} \tag{6.1}$$

6.2.1 The almost-sure upper bound

The earlier work of Harris and Williams [21] showed that there are *two* strictly-positive martingales Z_λ^- and Z_λ^+ defined as

$$Z_\lambda^\pm(t) := \sum_{k=1}^{N(t)} v_\lambda^\pm(Y_k(t)) e^{\lambda X_k(t) - E_\lambda^\pm t},$$

where v_λ^- and v_λ^+ are strictly positive eigenfunctions of the self-adjoint operator $\frac{1}{2}\lambda^2 A + R + Q_\theta$, with corresponding eigenvalues $E_\lambda^- < E_\lambda^+$. The explicit form for these eigenfunctions is

$$v_\lambda^\pm(y) = e^{\psi_\lambda^\pm y^2}$$

where $\psi_\lambda^\pm := \frac{1}{4} \pm \frac{\mu_\lambda}{2\theta}$, for a positive parameter μ_λ, and $\psi_\lambda^\pm$ are both positive for all $\lambda \in (\lambda_{\min}, 0)$.

A common theme is to overestimate indicator functions by an exponential, since it is often the case that this will bring in one of the martingales of the model: for $\lambda \in (\lambda_{\min}, 0)$,

$$\begin{aligned}\sum_{k=1}^{N(t)} \mathbf{1}\{X_k(t) \le -\gamma t, Y_k(t)^2 \ge \kappa^2 t\} &\le \sum_{k=1}^{N(t)} \exp\{\psi_\lambda^+(Y_k(t)^2 - \kappa^2 t)\} \exp\{\lambda(X_k(t) + \gamma t)\} \\ &= e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{(E_\lambda^- + \lambda\gamma - \kappa^2 \psi_\lambda^+)t}. \end{aligned} \tag{6.2}$$

(Importantly for this, the parameter ψ_λ^+ is positive and λ is negative; the functions c_λ^- and c_λ^+ are defined as $c_\lambda^\pm := E_\lambda^\pm/(-\lambda)$.)

The expression for $\Delta(\gamma, \kappa)$ as a Legendre-conjugate – see (6.1) – explains why $\Delta(\gamma, \kappa)$ relates to (6.2) above: by choosing λ at the infimum we get

$$N_t(\gamma, \kappa) \le e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) e^{\Delta(\gamma,\kappa)t}. \tag{6.3}$$

We remember that $N_t(\gamma, \kappa)$ takes only integer values, and a separate theorem by Harris and Git states that

$$\limsup_{t\to\infty} e^{-\lambda(c_\lambda^+ - c_\lambda^-)t} Z_\lambda^+(t) \le 0, \qquad \text{for each } \lambda \in (\lambda_{\min}, 0). \tag{6.4}$$

Thus if $\Delta(\gamma, \kappa) < 0$ we deduce that almost surely

$$N_t(\gamma, \kappa) = 0, \quad \text{eventually,}$$

whence $\lim_{t\to\infty} t^{-1} \log N_t(\gamma, \kappa) = -\infty$, as required.

On the other hand, if $\Delta(\gamma, \kappa) \ge 0$, (6.3) and (6.4) immediately imply that

$$\lim_{t\to\infty} t^{-1} \log N_t(\gamma, \kappa) \le \Delta(\gamma, \kappa).$$

□

6.2.2 A two-phase mechanism for the lower bound

For their proof of the almost-sure lower bound of Theorem 6.2.1, Harris and Git propose an explicit mechanism by which a sufficient number of particles will obtain a position near $(\gamma T, \kappa\sqrt{T})$ in the type-space domain at large times T. It is made up of two phases:

the long tread: Over a long period $[0, t]$, taking up nearly all of the time, a large number of particles will drift spatially with speed $\gamma\theta/(\theta + \kappa^2)$ – as if their type has had a modified occupation measure, as described by Harris and Williams [21];

the short climb: Following this, over a short period of time $[t, t+\tau]$ with τ a fixed small time ($\tau \ll t$), each of the particles from this group will have a small probability of further rushing to the large type position $\kappa\sqrt{t}$ whilst additionally gaining $\{\gamma\kappa^2/(\theta+\kappa^2)\}t$ in spatial position.

They have shown that the combination of these two phases will present us with sufficiently many particles at the space-type position $(\gamma T, \kappa\sqrt{T})$ at the large time $T = t + \tau$, concluding their proof of Theorem 6.2.1 – we refer the reader to their work for further details. Theorem 6.1.1 that we are going to prove makes up the *short climb* phase.

6.3 Large-deviations heuristics

We suppose that the number $\tau > 0$ is given and fixed; all our large-deviations results in this chapter will be considered as occurring over the fixed time-interval $[0, \tau]$. We here present some arguments concerning the large-deviations behaviour of the branching diffusion which the reader should take just as the *intuition* behind our later *rigourous* proofs.

By definition, under the measure $\tilde{P}$ $(= \tilde{P}^{0,0})$ the *spine* (ξ_s, η_s) satisfies

$$d\eta_s = \sqrt{\theta}dB_s - \frac{\theta}{2}\eta_s ds, \quad \text{and} \quad d\xi_s = \sqrt{a}\eta_s\, dW_s,$$

for two independent $\tilde{P}$-Brownian motions B_s and W_s. In the previous chapter we used the re-scaling of Definition 5.1.1 to effectively give the Brownian motion a variance of $1/\sqrt{T}$, suitable for a large-deviations analysis. Here the scaling is slightly different since we are always considering processes over the time-interval $[0, \tau]$; observing that for any $t > 0$,

$$d\left[\frac{\eta_s}{\sqrt{t}}\right] = \sqrt{\theta}\left[\frac{dB_s}{\sqrt{t}}\right] - \frac{\theta}{2}\left[\frac{\eta_s}{\sqrt{t}}\right]ds, \quad \text{and} \quad d\left[\frac{\xi_s}{t}\right] = \sqrt{a}\left[\frac{\eta_s}{\sqrt{t}}\right]\left[\frac{dW_s}{\sqrt{t}}\right],$$

we see that it is appropriate to work with the *re-scaled* spine $(\xi_s/t, \eta_s/\sqrt{t})$ since in this way we obtain a variance coefficient of $1/\sqrt{t}$ on the driving Brownian motions.

Definition 6.3.1 *For each $t > 0$ we define*

$$\xi_s^t := \xi_s/t, \quad \textit{and} \quad \eta_s^t := \eta_s/\sqrt{t},$$

and call (ξ_s^t, η_s^t) the ***re-scaled spine****. We note that under $\tilde{P}$ we have for $s \in [0, \tau]$:*

$$d\eta_s^t = \frac{\sqrt{\theta}}{\sqrt{t}}dB_s - \frac{\theta}{2}\eta_s^t ds, \quad \textit{and} \quad d\xi_s^t = \frac{\sqrt{a}\eta_s^t}{\sqrt{t}}dW_s,$$

for two independent $\tilde{P}$-Brownian motions B_s and W_s.

Suppose that we are given two paths: a type-path $y : [0, \tau] \to \mathbb{R}$ and a spatial-path $x : [0, \tau] \to \mathbb{R}$. Standard theory for a two-dimensional diffusion will deal with the large-deviations of the re-scaled spine (ξ_s^t, η_s^t), and on a *heuristic* level it states that the probability of the type-diffusion

η_s^t closely following y and the space-diffusion ξ_s^t closely following x is roughly

$$\exp\Big(-\frac{t}{2\theta}\int_0^\tau\big(\dot{y}_s+\frac{\theta}{2}y_s\big)^2\,\mathrm{d}s-\frac{t}{2}\int_0^\tau\frac{\dot{x}_s^2}{ay_s^2}\,\mathrm{d}s\Big),\tag{6.5}$$

for large enough t.

Given our work on the large-deviations principle for branching Brownian motion, one would make the reasonable guess that the probability that *at least one* of the *re-scaled* branching particles $\big(X_u(s)/t, Y_u(s)/\sqrt{t}\big)$ follows the type-path y_s and space-path x_s closely over the time interval $[0,\tau]$ is roughly

$$\exp\Big\{-\sup_{w\in[0,\tau]}\Big[\Big(\int_0^w\frac{1}{2\theta}\big(\dot{y}_s+\frac{\theta}{2}y_s\big)^2+\frac{1}{2}\frac{\dot{x}_s^2}{ay_s^2}-ry_s^2\,\mathrm{d}s\Big)t-\rho w\Big]\Big\},$$

when t is large. By standard optimization arguments (see Harris and Git [25] for example) this implies that the probability of at least one of the re-scaled branching particles being near the space-type position (β,κ) at a fixed time τ (which is also the event that the non-rescaled particles arrive near $(\beta t,\kappa\sqrt{t})$ of course) should be roughly

$$\exp\Big\{-\inf_{x,y}\sup_{w\in[0,\tau]}\Big[\Big(\int_0^w\frac{1}{2\theta}\big(\dot{y}_s+\frac{\theta}{2}y_s\big)^2+\frac{1}{2}\frac{\dot{x}_s^2}{ay_s^2}-ry_s^2\,\mathrm{d}s\Big)t-\rho w\Big]\Big\},\tag{6.6}$$

when t is large, and where the infimum is taken over all paths $x,y\in C[0,\tau]$ satisfying

$$y(0)=0, y(\tau)=\kappa, x(0)=0, x(\tau)=\beta.\tag{6.7}$$

This is typical in a large-deviations setting: although there are many possible trajectories that the (re-scaled) particles could travel along to get to a position (β,κ), the *dominant number* will have followed *optimal* paths.

Although the preceding arguments have been presented as if one is free to choose *any* paths x and y, we note that if $y_s=0$ when $\dot{x}_s\neq 0$ then a rigourous approach to these arguments may have problems with the term $\frac{\dot{x}_s^2}{ay_s^2}$ in (6.6) – the heuristics are not really problematic if we interpret this as saying that the probability is $e^{-\infty}=0$. On an intuitive level this is an expression of the fact that to have $y_s=0$ equates to turning off the Brownian variance in the spatial diffusion which in turn would imply that no spatial progress is possible and therefore $\dot{x}_s=0$ should be expected.

Harris and Git [25] state that for any given type-path y, the optimal space-path x for (6.6) under the constraint $x(\tau)=\beta$ will always be given by

$$x_s=\lambda\int_0^s ay_w^2\,\mathrm{d}w,\qquad\text{for } s\in[0,\tau],\tag{6.8}$$

for some value $\lambda\in\mathbb{R}$. Briefly, their arguments rely on the fact that in the definition of our model the spatial diffusion $X_u(s)$ of the branching particles can be seen as a time-changed Brownian motion where the time scaling is determined by its type process $Y_u(s)$:

$$X_u(s)=\hat{B}\Big(\int_0^s aY_u(w)^2\,\mathrm{d}w\Big)$$

for a Brownian motion $\hat{B}(\cdot)$. A measure change that introduces a linear drift of λ to this Brownian motion will give

$$X_u(s) = \tilde{B}\Big(\int_0^s aY_u(w)^2 \, \mathrm{d}w\Big) + \lambda \int_0^s aY_u(w)^2 \, \mathrm{d}w,$$

where $\tilde{B}(\cdot)$ is a Brownian motion under the new measure – this clearly relates to (6.8). Linear drifts are the optimal path (in a large-deviations sense) for a Brownian motion to be at a given point at a given time, and the constraint $x(\tau) = \beta$ for our problem will determine the value of λ in terms of the type-path y:

$$\lambda = \frac{\beta}{a\int_0^\tau y_s^2 \, \mathrm{d}s}. \tag{6.9}$$

Thus for the event being considered in Theorem 6.1.1, the spatial-path x is determined uniquely by (6.8) together with (6.9). Therefore an equivalent statement of our large-deviations is that the probability of at least one of the re-scaled branching particles being near the space-type position (β, κ) at a fixed time τ is roughly

$$\exp\Big\{ -\inf_{\lambda, y} \sup_{w\in[0,\tau]} \Big[\Big(\int_0^w \frac{1}{2\theta}\big(\dot{y}_s + \frac{\theta}{2} y_s\big)^2 + \frac{a\lambda^2}{2} y_s^2 - r y_s^2 \, \mathrm{d}s \Big) t - \rho w \Big] \Big\}, \tag{6.10}$$

when t is large, and where the infimum is taken over all paths $y \in C[0,\tau]$ and all $\lambda \in (\lambda_{\min}, 0)$ satisfying

$$y(0) = 0, y(\tau) = \kappa, \quad \lambda := \frac{\beta}{a\int_0^\tau y_s^2 \, \mathrm{d}s}. \tag{6.11}$$

Harris and Git [25] presented alternative heuristic arguments based on birth-death processes to arrive at the expression (6.10). Then using Euler-Lagrange techniques they showed that the specific path

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad s \in [0,\tau] \tag{6.12}$$

is optimal for this expression, where

$$\mu := \mu_\lambda = \frac{\sqrt{\theta(\theta - 8r - 4a\lambda^2)}}{2}, \tag{6.13}$$

and $\lambda \in (\lambda_{\min}, 0)$ is dependent on the choice of τ (which we are anyway considering as fixed throughout) and is chosen to satisfy

$$\frac{\beta}{a\lambda} = \kappa^2 \Big(\frac{\coth \mu_\lambda \tau}{\mu_\lambda} - \frac{\tau}{2\sinh^2 \mu_\lambda \tau} \Big). \tag{6.14}$$

We refer the reader to Harris and Git [25] for details of these relationships between the parameters, but note that particles staying close to this path will arrive near $y(\tau) = \kappa$ at time τ in agreement with the heuristics.

As we mentioned just before the statement of Theorem 6.1.1, our spine techniques will naturally use the path y_s defined at (6.12) together with x_s defined at (6.8), since they are the optimal paths (in a large-deviations sense) for accumulating particles near the point $(\beta t, \kappa\sqrt{t})$ at time τ. Our spine proof of Theorem 6.1.1 will result in a proof of the following stronger result, from which Theorem 6.1.1 would actually follow as a corollary.

Theorem 6.3.2 *Let $\tau > 0$ be fixed. For any $\kappa \in \mathbb{R}$ and $\lambda \in (\lambda_{\min}, 0)$ we define two continuous paths on $[0,\tau]$,*

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad x_s := a\lambda \int_0^s y_w^2 \,\mathrm{d}w, \qquad s \in [0,\tau], \tag{6.15}$$

and note that at time τ these paths reach the points $y_\tau = \kappa$ and $x_\tau = \beta$ where

$$\beta := a\lambda \int_0^\tau y_w^2 \,\mathrm{d}w = \kappa^2 \Big(\frac{\coth \mu_\lambda \tau}{2\mu_\lambda} - \frac{\tau}{2\sinh^2 \mu_\lambda \tau} \Big).$$

Then, for large t, the probability that at least one of the typed branching particles $(X_u(s), Y_u(s))$ will stay near $(tx_s, \sqrt{t}y_s)$ throughout $s \in [0,\tau]$ (given that the original ancestor was at the space-type origin) has a large-deviations lower bound: for all $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \forall s \in [0,\tau], \big|X_u(s) - tx_s\big| < \delta t, \big|Y_u(s) - \sqrt{t}y_s\big| < \delta'\sqrt{t}\Big) \geq -\Theta(\beta,\kappa).$$

The fact that our spine proofs can give such good path results, together with the work of the previous chapter for BBM, suggest that it will be possible to develop the ideas that we use in this chapter to proofs of large-deviations principles for many branching-diffusion models (see Harris and Git [25] for a discussion of such principles). This consistent spine approach can be applied to achieve excellent results for branching diffusions essentially because the Many-to-One theorem 3.6.5 and martingale estimates with the spine decomposition 3.5.1 will always reduce the branching problem to an issue of large-deviations for a single diffusing particle (the spine) which are already well studied.

6.4 Martingales and measures

Although we have already indicated a specific path at (6.12), it should be noted that in our proofs we use properties of this path only at a few points – elsewhere the techniques can be applied in general to any path. Therefore the reader may suppose that $y : [0,\tau] \to \mathbb{R}$ is any given and fixed path, and we shall be very careful to highlight those points where we use specific properties of the path defined at (6.12). Also, to keep notational complexity to a reasonable minimum we tend not to make the dependencies of the martingales and action functionals on the underlying chosen paths explicit in the notation.

Estimates on a martingale can give us a lower-bound for the large-deviations events of our branching diffusion, and the expressions that we considered in the heuristics of the previous section can be taken as a starting point. For any $t > 0$ and any given $y : [0,\tau] \to \mathbb{R}$ that is square-integrable along with its derivative

$$\exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}} \int_0^w (\dot{y}_s + \frac{\theta}{2} y_s) \,\mathrm{d}B_s - \frac{t}{2\theta} \int_0^w (\dot{y}_s + \frac{\theta}{2} y_s)^2 \,\mathrm{d}s \Big) \times \exp\Big(\sqrt{at}\lambda \int_0^w y_s \,\mathrm{d}W_s - \frac{a\lambda^2 t}{2} \int_0^w y_s^2 \,\mathrm{d}s \Big),$$

is a strictly-positive $\tilde{P}$-martingale over the time period $w \in [0,\tau]$ (see Øksendal [46] for example). As one part of the change of measure defined below, this martingale will introduce drift *where we use $B_u(s)$ and $W_u(s)$ to denote the P-Brownian motions driving the type and spatial processes of particle u in the branching diffusion.*

The conditional-expectation relationship between $\tilde{\zeta}_t$ and $Z_t(w)$ means that Z_t is the Radon-Nikodym derivative between the branching-particle measures $\mathbb{Q}_t$ and P:

Theorem 6.4.3 *We define the measure $\mathbb{Q}_t$ on $(\tilde{T}, \mathcal{F}_\tau, (\mathcal{F}_w))$ as the restriction of the measure $\tilde{\mathbb{Q}}_t$ to the algebra $\mathcal{F}_\tau$: $\mathbb{Q}_t := \tilde{\mathbb{Q}}_t|_{\mathcal{F}_\tau}$. Using a proof similar to that used in Theorem 3.4.4, it can be shown that for each $w \in [0, \tau]$,*

$$\left.\frac{\mathrm{d}\mathbb{Q}_t}{\mathrm{d}P}\right|_{\mathcal{F}_w} = Z_t(w).$$

For our proof of Theorem 6.1.1 (and its stronger version of Theorem 6.3.2) it is important to know how quickly $Z_t(\tau)$ grows under the measure $\mathbb{Q}_t$. Since Z_t is a submartingale with respect to the measure $\mathbb{Q}_t$, we can use Doob's submartingale inequality together with the following result on the exponential growth rate of $\mathbb{Q}(Z_t(\tau)^\alpha)$ for $\alpha \in [0, 1]$.

Theorem 6.4.4 *For the specific path y defined at (6.12), and for any $\alpha \in [0, 1]$ we have*

$$\limsup_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t\big(Z_t(\tau)^\alpha\big) \leq \alpha J(\tau) + \alpha^2 M(\tau),$$

where we define

$$J(w) := \int_0^w \left[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2 - ry_s^2\right] \mathrm{d}s, \tag{6.20}$$

and

$$M(w) := \int_0^w \left[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2\right] \mathrm{d}s. \tag{6.21}$$

We emphasize that without the technology of spines the proof of this result would be exceptionally difficult – witness the proofs of $\mathcal{L}^p$-convergence in chapter 2 for the simpler martingale Z_λ, where the classical approach succeeded mainly thanks to the (ad-hoc) inequality of Lemma 2.5.2 and the fact that we were dealing with a finite-dimensional type space. It is notoriously difficult to deal with operations like $Z_t(\tau)^\alpha$ since these martingales Z_t are defined via *sums*, and classical inequalities will tend to not be good enough.

In contrast to this, the spine decomposition gives us a proper methodology for reducing the additive structure of these martingales to essentially a single-particle problem, and since it does this through a conditional-expectation operation rather than with an inequality, it is *exact* and therefore can lead to tight estimates that are useful. Due to its length, we dedicate the whole of section 6.6 to the spine proof of this above theorem, and now proceed to show how this result can be used to obtain the upper-bound on $Z_t(\tau)$ that we require for Theorem 6.1.1.

Given Theorem 6.4.4, we can use Doob's submartingale inequality to prove the following:

Theorem 6.4.5 *Let $\tau > 0$ be fixed. Then for all $\varepsilon > 0$,*

$$\lim_{t\to\infty} \mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) \leq e^{(J(\tau)+\varepsilon)t}\Big) \to 1.$$

terms into the diffusions η_s and ξ_s such that $\eta^t_s \sim y_s$ and $\xi^t_s \sim a\lambda y^2_s$ when t is large, and we note a comparison between this martingale and the expression (6.10) above.

The process n_w which counts the number of fission times on the spine up to time w is a Cox process of rate $R(\eta_s)$ and therefore

$$w \mapsto e^{-\int_0^w R(\eta_s)\,\mathrm{d}s} 2^{n_w}$$

is a $\tilde{P}$-martingale too (a similar martingale was used by Kyprianou [35] for BBM). We use the product of these two martingales to define a new measure:

Theorem 6.4.1 *For $t > 0$ we define a measure $\tilde{\mathbb{Q}}_t$ on the filtered space $(\tilde{\mathcal{T}}, \tilde{\mathcal{F}}_\tau, (\tilde{\mathcal{F}}_w)_{w\in[0,\tau]})$*

$$\left.\frac{\mathrm{d}\tilde{\mathbb{Q}}_t}{\mathrm{d}\tilde{P}}\right|_{\tilde{\mathcal{F}}_w} := e^{-\int_0^w R(\eta_s)\,\mathrm{d}s} 2^{n_w} \times \exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (\dot{y}_s + \frac{\theta}{2} y_s)\,\mathrm{d}B_s - \frac{t}{2\theta}\int_0^w (\dot{y}_s + \frac{\theta}{2} y_s)^2\,\mathrm{d}s\Big)$$
$$\times \exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_s - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big). \quad (6.16)$$

We give this $\tilde{\mathcal{F}}_w$-martingale the name $\tilde{\zeta}(w) := \mathrm{d}\tilde{\mathbb{Q}}_t/\mathrm{d}\tilde{P}|_{\tilde{\mathcal{F}}_w}$. Under the measure $\tilde{\mathbb{Q}}_t$ we can give a pathwise construction of the branching-diffusion $(\mathbb{X}_s)_{s\in[0,\tau]}$:

- *the spine process (ξ_s, η_s) starts at $(0,0)$ and diffuses as a solution to*

$$\mathrm{d}(\eta^t_s - y_s) = \frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}\tilde{B}_s - \frac{\theta}{2}(\eta^t_s - y_s)\mathrm{d}s \quad (6.17)$$

  and

$$\mathrm{d}\xi^t_s = \frac{\sqrt{a}}{\sqrt{t}}\eta^t_s\,\mathrm{d}\tilde{W}_s + a\lambda y_s \eta^t_s\,\mathrm{d}s, \quad (6.18)$$

  so that the type process η^t_s will be an $OU(\theta/t, \theta/2)$ along the path y, and the spatial motion ξ_s has a drift component added;

- *at the accelerated rate $2R(\eta_s)$ the spine undergoes fission producing two particles;*

- *with equal probability, one of these two particles is selected to continue the spine;*

- *the other particle initiates, from its birth space-type position, an independent copy of the original P branching diffusion with normal branching rate $R(\cdot)$.*

It is by now familiar to derive an additive martingale from (6.16) by projecting it onto the filtration of the branching diffusion particles, as laid out in Lemma 3.3.7.

Definition 6.4.2 *For each $t > 0$ we define a P-martingale (with respect to the filtration $(\mathcal{F}_w)_{0\le w\le\tau}$) as*

$$Z_t(w) = \tilde{P}\big(\tilde{\zeta}_t(w)|\mathcal{F}_w\big).$$

Using the representation (3.4) it can be shown that

$$Z_t(w) = \sum_{u\in N_w} e^{-r\int_0^w Y_u(s)^2\,\mathrm{d}s - \rho w} \exp\Big(\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w (\dot{y}_s + \frac{\theta}{2} y_s)\,\mathrm{d}B_u(s) - \frac{t}{2\theta}\int_0^w (\dot{y}_s + \frac{\theta}{2} y_s)^2\,\mathrm{d}s\Big)$$
$$\times \exp\Big(\sqrt{at}\lambda\int_0^w y_s\,\mathrm{d}W_u(s) - \frac{a\lambda^2 t}{2}\int_0^w y_s^2\,\mathrm{d}s\Big), \quad (6.19)$$

Proof: For a given $\varepsilon > 0$ and for any $\alpha \in [0,1]$, Doob's inequality combines with our estimate of $\mathbb{Q}_t(Z_t(\tau)^\alpha)$ to give

$$\mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t}\Big) = \mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s)^\alpha > e^{\alpha(J(\tau)+\varepsilon)t}\Big) \le \frac{\mathbb{Q}_t\big(Z_t(\tau)^\alpha\big)}{e^{\alpha(J(\tau)+\varepsilon)t}}.$$

From Theorem 6.4.4 we know that for each $\alpha \in [0,1]$ and for all large t we have

$$\mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t}\Big) \le e^{(\alpha M-\varepsilon)\alpha t}.$$

For any $\alpha \in (0, \varepsilon/M)$ this is a decaying exponential and so it follows that

$$\lim_{t\to\infty} \mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) > e^{(J(\tau)+\varepsilon)t}\Big) \to 0.$$

□

For the specific y defined at (6.12), it can be shown that

$$J(\tau) = \int_0^\tau \Big[\frac{1}{2\theta}(\dot{y}_s + \frac{\theta}{2}y_s)^2 + \frac{a\lambda^2}{2}y_s^2 - ry_s^2\Big]\,\mathrm{d}s = \lambda\beta + \kappa^2\Big(\frac{1}{4} + \frac{\mu_\lambda}{2\theta}\coth\mu_\lambda\tau\Big),$$

where we recall that this $\lambda \in (\lambda_{\min}, 0)$ was specifically determined by (6.14). In fact, Harris and Git [25] explain that this choice of λ was optimal in that

$$\lambda\beta + \kappa^2\Big(\frac{1}{4} + \frac{\mu_\lambda}{2\theta}\coth\mu_\lambda\tau\Big) = \sup_\gamma\Big\{\gamma\beta + \kappa^2\Big(\frac{1}{4} + \frac{\mu_\gamma}{2\theta}\coth\mu_\gamma\tau\Big)\Big\}. \tag{6.22}$$

On the other hand we can find a similar representation for the parameter $\Theta(\beta,\kappa)$: if we define

$$\bar{\lambda} := \sqrt{\frac{\beta^2\theta(\theta-8r)}{a^2\kappa^4 + 4a\theta\beta^2}}, \quad \text{so that } \mu_{\bar{\lambda}} = \frac{\kappa^2\sqrt{\theta(\theta-8r)}}{2\sqrt{\kappa^4 + 4\theta\beta^2/a}},$$

then

$$\Theta(\beta,\kappa) = \bar{\lambda}\beta + \kappa^2\psi^+_{\bar{\lambda}} = \lim_{\tau\to\infty}\sup_\gamma\Big\{\gamma\beta + \kappa^2\Big(\frac{1}{4} + \frac{\mu_\gamma}{2\theta}\coth\mu_\gamma\tau\Big)\Big\},$$

where we recall that

$$\psi^+_{\bar{\lambda}} := \frac{1}{4} + \frac{\mu_{\bar{\lambda}}}{2\theta}.$$

In this way it can be deduced from (6.22) that

$$J(\tau) \uparrow \Theta(\beta,\kappa),$$

which in turn implies the following corollary to Theorem 6.4.5:

Corollary 6.4.6 *Let $\tau > 0$ be fixed. Then for all $\varepsilon > 0$,*

$$\lim_{t\to\infty} \mathbb{Q}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) \le e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big) \to 1.$$

6.5 Proving the large-deviations lower bound

Barring the proof of Theorem 6.4.4 which we cover fully in section 6.6, we now have all the information required to prove the large-deviations lower-bound for the short-climb event of Theorem 6.1.1, which we recall as stating that for large t, the probability that at least one of the branching particles will be near $(\beta t, \kappa\sqrt{t})$ at time τ (where $\tau, \beta, \kappa \in \mathbb{R}$ with $\beta < 0$ and $\tau > 0$ are given and fixed) has a large-deviations lower bound: for all $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau : \big|X_u(\tau) - \beta t\big| < \delta t, \big|Y_u(\tau) - \kappa\sqrt{t}\big| < \delta'\sqrt{t}\Big) \geq -\Theta(\beta, \kappa).$$

Noting that this event is $\mathcal{F}_\tau$-measurable since it depends only on the branching particles and does not refer to the spine, it follows that on this event the change of measure is carried out by Z_t, as noted in Theorem 6.4.3. The upper bound that we have derived for Z_t at Corollary 6.4.6 will serve as a lower bound for $1/Z_t(\tau)$ in this change of measure, and will combine with the fact that under the measure $\tilde{\mathbb{Q}}_t$ (for large t) we know that the spine will carry out the large-deviations behaviour that we want.

Throughout this proof we are focussing on the specific path

$$y_s := \kappa \frac{\sinh \mu_\lambda s}{\sinh \mu_\lambda \tau}, \quad s \in [0, \tau]$$

where $\lambda \in (\lambda_{\min}, 0)$ satisfies

$$\frac{\beta}{a\lambda} = \kappa^2 \Big(\frac{\coth \mu_\lambda \tau}{\mu_\lambda} - \frac{\tau}{2\sinh^2 \mu_\lambda \tau} \Big).$$

as discussed at (6.14). We define the event that the space-type location $\big(X_u(s), Y_u(s)\big)$ of a particular particle u remains near $(a\lambda t \int_0^s y_w^2 \, \mathrm{d}w, \sqrt{t}y_s)$ throughout the interval $s \in [0, \tau]$:

$$A_t(u) := \Big(\forall s \in [0, \tau], \big|X_u(s) - a\lambda t \int_0^s y_w^2 \, \mathrm{d}w\big| < \delta, \big|Y_u(s) - \sqrt{t}y_s\big| < \delta' \Big),$$

where $\delta, \delta' > 0$ are given and fixed. Then for any $\varepsilon > 0$,

$$\begin{aligned}
P_t\Big(\exists u \in N_\tau \text{ such that } A_t(u)\Big) &= \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u)\Big) \\
&\geq \mathbb{Q}_t\Big(\frac{1}{Z_t(\tau)}; \exists u \in N_\tau, A_t(u); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big) \\
&\geq e^{-(\Theta(\beta,\kappa)+\varepsilon)t}\, \mathbb{Q}_t\Big(\exists u \in N_\tau, A_t(u); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big) \\
&\geq e^{-(\Theta(\beta,\kappa)+\varepsilon)t}\, \tilde{\mathbb{Q}}_t\Big(A_t(\xi); \sup_{s\in[0,\tau]} Z_t(s) \leq e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big). \qquad (6.23)
\end{aligned}$$

Given (6.17) and (6.18), standard theory says that under the measure $\tilde{\mathbb{Q}}_t$ (with t large) the re-scaled spine (ξ_s^t, η_s^t) will tend to stay close to the space-type paths $(a\lambda \int_0^s y_w^2 \, \mathrm{d}w, y_s)$ over the whole time interval $[0, \tau]$:

$$\xi_s^t \sim a\lambda \int_0^s y_w^2 \, \mathrm{d}w, \quad \text{and} \quad \eta_s^t \sim y_s,$$

by which we mean that for a fixed $\tau > 0$ and any $\delta, \delta' > 0$,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\big|\xi^t_s - a\lambda \int_0^s y_w^2 \,\mathrm{d}w\big| < \delta, \big|\eta^t_s - y_s\big| < \delta', \text{for all } s \in [0,\tau]\Big) \to 1,$$

which can equally be written as:

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\big|\xi_s - a\lambda t \int_0^s y_w^2 \,\mathrm{d}w\big| < \delta t, \big|\eta_s - y_s\sqrt{t}\big| < \delta'\sqrt{t}, \text{for all } s \in [0,\tau]\Big) \to 1,$$

which is exactly the statement that

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t(A_t(\xi)) = 1.$$

Corollary 6.4.6 says,

$$\lim_{t\to\infty} \tilde{\mathbb{Q}}_t\Big(\sup_{s\in[0,\tau]} Z_t(s) \le e^{(\Theta(\beta,\kappa)+\varepsilon)t}\Big) = 1,$$

and since $\varepsilon > 0$ was arbitrary, it follows from (6.23) that for all $\delta, \delta' > 0$,

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u, \forall s \in [0,\tau], \big|X_u(s) - a\lambda t \int_0^s y_w^2 \,\mathrm{d}w\big| < \delta, \big|Y_u(s) - \sqrt{t}y_s\big| < \delta'\Big) \ge -\Theta(\beta,\kappa),$$

which gives the proof of the stronger version at Theorem 6.3.2. The constraints of (6.11) state that for the particular y and λ we chose above, we have

$$\sqrt{t}y(\tau) = \kappa\sqrt{t}, \quad \text{and} \quad a\lambda t \int_0^\tau y_s^2 \,\mathrm{d}w = \beta t,$$

and therefore we can also deduce the weaker result to complete our proof of Theorem 6.1.1:

$$\liminf_{t\to\infty} t^{-1} \log P\Big(\exists u \in N_\tau, \big|X_u(\tau) - \beta t\big| < \delta, \big|Y_u(\tau) - \kappa\sqrt{t}\big| < \delta'\Big) \ge -\Theta(\beta,\kappa).$$

This completes the proof of the short-climb large-deviations lower bound. $\square$

6.6 A spine proof of the martingale upper-bound

In this section we use the spine decomposition of the martingale Z_t to prove Theorem 6.4.4. It is Jensen's inequality that immediately allows us to concentrate on the spine decomposition:

$$\mathbb{Q}_t\big(Z_t(\tau)^\alpha\big) \le \tilde{\mathbb{Q}}_t\Big(\tilde{\mathbb{Q}}_t\big(Z_t(\tau)|\tilde{\mathcal{G}}_\infty\big)^\alpha\Big), \qquad \text{since } \alpha \in [0,1].$$

The spine decomposition of $Z_t(\tau)$ is

$$\begin{aligned}\tilde{\mathbb{Q}}_t\big(Z_t(\tau)|\tilde{\mathcal{G}}_\infty\big) &= e^{-r\int_0^\tau \eta_s^2 \,\mathrm{d}s - \rho\tau} e^{[\sqrt{a}\lambda \int_0^\tau y_s \,\mathrm{d}W_s - \frac{a\lambda^2}{2}\int_0^\tau y_s^2 \,\mathrm{d}s] + [\frac{1}{\sqrt{\theta}}\int_0^\tau (\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B_s - \frac{1}{2\theta}\int_0^\tau (\dot{y}_s + \frac{\theta}{2}y_s)^2 \,\mathrm{d}s]} \\ &+ \sum_{k=1}^{n_\tau} e^{-r\int_0^{S_k} \eta_s^2 \,\mathrm{d}s - \rho S_k} e^{[\sqrt{a}\lambda \int_0^{S_k} y_s \,\mathrm{d}W_s - \frac{a\lambda^2}{2}\int_0^{S_k} y_s^2 \,\mathrm{d}s] + [\frac{1}{\sqrt{\theta}}\int_0^{S_k} (\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}B_s - \frac{1}{2\theta}\int_0^{S_k} (\dot{y}_s + \frac{\theta}{2}y_s)^2 \,\mathrm{d}s]}.\end{aligned}$$

We consider the two parts of this spine decomposition separately – the **spine term** and then the **sum term** – and aim to show that they both have exponential growth of the same order.

Definition 6.6.1 *We define*

$$\textbf{spine term} := e^{-r\int_0^\tau \eta_s^2\,\mathrm{d}s-\rho\tau}e^{[\sqrt{a}\lambda\int_0^\tau y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^\tau y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]},$$

and

$$\textbf{sum term} :=$$

$$\sum_{k=1}^{n_\tau} e^{-r\int_0^{S_k}\eta_s^2\,\mathrm{d}s-\rho S_k}e^{[\sqrt{a}\lambda\int_0^{S_k} y_s\,\mathrm{d}W_s-\frac{a\lambda^2}{2}\int_0^{S_k} y_s^2\,\mathrm{d}s]+[\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}B_s-\frac{1}{2\theta}\int_0^{S_k}(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s]}.$$

In each case we first use some martingale techniques to factor out exponential terms that give us the correct growth rate (and here we are guided by the heuristics), and then use Varadhan's lemma to show that the remaining terms do not contribute any further exponential growth. The spine term is simpler to deal with and is considered first.

Factoring out the spine term

Girsanov's theorem (see Øksendal [46]) states that under the new measure $\tilde{\mathbb{Q}}_t$ we have

$$\mathrm{d}B_s = \mathrm{d}\tilde{B}_s + \frac{\sqrt{t}}{\sqrt{\theta}}\Big(\dot{y}_s + \frac{\theta}{2}y_s\Big)\mathrm{d}s, \quad \text{and} \quad \mathrm{d}W_s = \mathrm{d}\tilde{W}_s + \sqrt{at}\lambda y_s\,\mathrm{d}s, \tag{6.24}$$

which can both be substituted into the spine term to give,

$$\begin{aligned}\textbf{spine term} &= e^{t\int_0^\tau \frac{1}{2\theta}(\dot{y}_s+\frac{\theta}{2}y_s)^2+\frac{a\lambda^2}{2}y_s^2\,\mathrm{d}s-\rho\tau} \times e^{-r\int_0^\tau \eta_s^2\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]} \\ &= e^{tJ(\tau)-\rho\tau} \times e^{rt\int_0^\tau[(\eta_s^t)^2-y_s^2]\,\mathrm{d}s}\, e^{[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.\end{aligned} \tag{6.25}$$

Using the standard martingale

$$e^{\alpha\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s-\alpha^2\frac{a\lambda^2 t}{2}\int_0^\tau y_s^2\,\mathrm{d}s},$$

we can factor out one of the terms of the expectation:

$$\begin{aligned}\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) &= e^{\alpha tJ(\tau)-\alpha\rho\tau}\,\tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\, e^{\alpha[\sqrt{at}\lambda\int_0^\tau y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big) \\ &= e^{\alpha tJ(\tau)-\alpha\rho\tau}e^{\alpha^2\frac{a\lambda^2 t}{2}\int_0^\tau y_s^2\,\mathrm{d}s}\,\tilde{\mathbb{Q}}_t\Big(e^{\alpha r\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\, e^{\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big).\end{aligned}$$

This final expectation can be dealt with by another change of measure:

$$\begin{aligned}&\tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}e^{\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\Big) \\ &= e^{\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s} \times \tilde{\mathbb{Q}}_t\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}e^{\frac{\alpha\sqrt{t}}{\sqrt{\theta}}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s-\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\Big), \\ &= e^{\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\Big),\end{aligned}$$

where we have used the martingale

$$e^{\frac{\alpha\sqrt{t}}{\sqrt{\theta}}\int_0^\tau \dot{y}_s+\frac{\theta}{2}y_s\,\mathrm{d}\tilde{B}_s-\frac{\alpha^2 t}{2\theta}\int_0^\tau(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}$$

to change the measure from $\tilde{\mathbb{Q}}_t$ to $\tilde{\mathbb{Q}}_t^\alpha$. Another application of the Girsanov theorem implies that under the measure $\tilde{\mathbb{Q}}_t^\alpha$, the re-scaled process η_s^t satisfies (where $\bar{B}_s$ is a Brownian motion)

$$\mathrm{d}\big(\eta_s^t - (1+\alpha)y\big) = \frac{\sqrt{\theta}}{\sqrt{t}}\mathrm{d}\bar{B}_s - \frac{\theta}{2}\big(\eta_s^t - (1+\alpha)y_s\big)\mathrm{d}s \tag{6.26}$$

which is to say that η^t is an $\mathrm{OU}(\frac{\theta}{t}, \frac{\theta}{2})$ along the *perturbed* path $(1+\alpha)y$.

Putting this all together we are left with a neat factorization expressed in terms of the re-scaled type process η_s^t:

$$\begin{aligned}\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) &= e^{\alpha t J(\tau) - \alpha\rho\tau} e^{\alpha^2 t M} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big), \\ &\leq e^{\alpha t J(\tau)} e^{\alpha^2 t M} \times \tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big)\end{aligned} \tag{6.27}$$

The term $\alpha\rho\tau$ becomes insignificant in the large deviations limit (for which $t \to \infty$), and therefore it is convenient to remove it here.

The martingale techniques have now played their part, and we move on to use Varadhan's lemma to show that the term $\tilde{\mathbb{Q}}_t^\alpha\big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\big)$ decays exponentially as $t \to \infty$.

A first application of Varadhan's lemma

Under the measure $\tilde{\mathbb{Q}}_t^\alpha$ the process η^t is an $\mathrm{OU}(\frac{\theta}{t}, \frac{\theta}{2})$ along the perturbed path $(1+\alpha)y$ (or equivalently we can say that $[\eta_s^t - (1+\alpha)y_s]$ is an $\mathrm{OU}(\frac{\theta}{t}, \frac{\theta}{2})$), and therefore it satisfies a large-deviations principle:

Theorem 6.6.2 *If we use the notation η^t to refer to the element (path) in $C[0,\tau]$ defined by*

$$\eta^t(s) := \eta_s^t, \quad \textit{for } s \in [0,\tau]$$

then there is a large-deviations principle for η^t with respect to the measure $\tilde{\mathbb{Q}}_t^\alpha$:

- *Upper bound: If C is a closed subset of $C[0,\tau]$ then*

$$\limsup_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha(\eta_s^t \in C) \leq -\inf_{g\in C} I(g,\tau),$$

- *Lower bound: If V is an open subset of $C[0,\tau]$ then*

$$\liminf_{t\to\infty} t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha(\eta_s^t \in V) \geq -\inf_{g\in V} I(g,\tau),$$

where

$$I(g,w) := \int_0^w \frac{1}{2\theta}\Big[\dot{g}_s + \frac{\theta}{2}g_s - (1+\alpha)\big(\dot{y}_s + \frac{\theta}{2}y_s\big)\Big]^2\,\mathrm{d}s.$$

if $g \in C[0,\tau]$ with $g(0) = 0$ is square-integrable along with its derivative; otherwise we define $I(g) = \infty$.

Given the upper bound (6.27) we now want to understand the behaviour of the expectation term $\tilde{\mathbb{Q}}_t^\alpha\Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\Big)$ for large t. Varadhan's lemma is a common way to deal with expectations of this form, and we quote the following from Dembo and Zeitouni [12].

Theorem 6.6.3 (Varadhan) *Let $(X^t)_{t\geq 0}$ be a family of random variables taking values in the space $\mathcal{X}$, and let μ_t denote the probability measures associated with $(X^t)_{t\geq 0}$.*

Suppose that the measures μ_t satisfy the LDP with a good rate function $I : \mathcal{X} \to [0,\infty]$, and let $\phi : \mathcal{X} \to \mathbb{R}$ be any continuous function. Assume further that the following moment condition holds for some $\gamma > 1$,

$$\limsup_{t\to\infty} t^{-1} \log \mathbb{E}\Big[e^{\gamma t \phi(X^t)}\Big] < \infty. \tag{6.28}$$

Then

$$\lim_{t\to\infty} t^{-1} \log \mathbb{E}\Big[e^{t\phi(X^t)}\Big] = \sup_{x\in\mathcal{X}} \big[\phi(x) - I(x)\big].$$

This powerful result will confirm our hopes that the exponential will decay as $t \to \infty$.

Theorem 6.6.4 *For each $\alpha > 0$ the expectation decays exponentially to 0:*

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha \big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\big) < 0. \tag{6.29}$$

For small α we can give more precise expression of the exponential decay:

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha \big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\big) = \\ -\alpha^2 \left\{ k_1 \left[\int_0^\tau r y_s^2 \,\mathrm{d}s\right] + k_2 \left[\frac{1}{2\theta} \int_0^\tau \big(\dot{y}_s + \frac{\theta}{2} y_s\big)^2 \,\mathrm{d}s\right]\right\} + o(\alpha^2), \quad \textit{as } \alpha \to 0,$$

where k_1, k_2 are strictly positive.

Proof: Given the large-deviations principle stated in Theorem 6.6.2, we shall be equating $\mathcal{X} = C[0,\tau]$, $X^t = \eta^t$ and $\mu_t = \tilde{\mathbb{Q}}_t^\alpha$ and have $\phi(\eta^t) = \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s$, and the moment condition (6.28) is satisfied because

$$\tilde{\mathbb{Q}}_t^\alpha \big(e^{2\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\big) < e^{2\alpha r t \int_0^\tau y_s^2\,\mathrm{d}s}.$$

Varadhan's lemma then says that

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha \big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\big) = \sup_{z\in C_0[0,\tau]} \left\{ \Big(\int_0^\tau \alpha r \big[y_s^2 - z_s^2\big]\,\mathrm{d}s\Big) - I(z,\tau)\right\}. \tag{6.30}$$

Standard Euler-Lagrange techniques for maximizing the right-hand integral lead to the following differential equation for z:

$$\ddot{z}_s - \Big(\frac{\theta^2}{4} + 2\theta\alpha r\Big) z_s = (1+\alpha)\ddot{y}_s - \frac{\theta^2}{4}(1+\alpha) y_s, \tag{6.31}$$

which in general will give the optimal path as a solution in terms of the given path y.

With the *specific* path (6.12) that resulted from the Harris and Git optimizations of the large-deviations heuristics, it is relatively simple to solve (6.31) and find that the optimal path z is just a constant multiple of the path y:

$$z_s = K_\alpha y_s, \qquad \text{where } K_\alpha := \frac{\mu_\lambda^2 - \theta^2/4}{\mu_\lambda^2 - \theta^2/4 - 2\theta\alpha r}(1+\alpha). \tag{6.32}$$

Substituting for z into (6.30) we find that

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha \left(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s}\right)$$

$$= \alpha(1 - K_\alpha^2)\left[\int_0^\tau r y_s^2 \,\mathrm{d}s\right] - (K_\alpha - (1+\alpha))^2 \left[\frac{1}{2\theta}\int_0^\tau \left(\dot{y}_s + \frac{\theta}{2} y_s\right)^2 \mathrm{d}s\right], \tag{6.33}$$

and the following simple bound on K_α implies that this is a negative quantity

Lemma 6.6.5 *For all $\alpha > 0$,*

$$1 < K_\alpha < 1 + \alpha. \tag{6.34}$$

This small lemma can be proved with simple algebra from the definition of μ_λ given at (6.13): we can use this to show that $\mu_\lambda^2 - \theta^2/4 = -2\theta r - a\theta\lambda^2 < 0$, from which it follows that

$$\frac{1}{1+\alpha} < \frac{\mu_\lambda^2 - \theta^2/4}{\mu_\lambda^2 - \theta^2/4 - 2\theta\alpha r} < 1.$$

If we make a Taylor expansion about $\alpha = 0$:

$$\frac{\mu_\lambda^2 - \theta^2/4}{\mu_\lambda^2 - \theta^2/4 - 2\theta\alpha r} = \frac{1}{1 - k\alpha} = 1 + k\alpha + k^2\alpha^2 + o(\alpha^2) + \cdots$$

where $k := \frac{2\theta r}{\mu_\lambda^2 - \theta^2/4}$, it follows that for strictly positive constants k_1 and k_2,

$$\alpha(1 - K_\alpha^2) = -k_1\alpha^2 + o(\alpha^2), \quad \text{and} \quad (K_\alpha - (1+\alpha))^2 = k_2\alpha^2 + o(\alpha^2) \qquad \text{as } \alpha \to 0,$$

completing the proof □

Dealing with the sum term

Focusing on the sum term, we can again substitute for $\mathrm{d}W_s$ and $\mathrm{d}B_s$ with (6.24) and immediately factor out the term $J(S_k)$ by over-estimating:

$$\begin{aligned}\textbf{sum term} &= \sum_{k=1}^{n_\tau} e^{tJ(S_k) - \rho S_k} e^{r\int_0^{S_k}[y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s} e^{[\sqrt{at}\lambda \int_0^{S_k} y_s \,\mathrm{d}\tilde{W}_s] + [\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}\bar{B}_s]} \\ &\le e^{t\left(\sup_{0\le w\le\tau} J(w)\right)} \sum_{k=1}^{n_\tau} e^{r\int_0^{S_k} \eta_s^2 - y_s^2 \,\mathrm{d}s} e^{[\sqrt{a}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s] + [\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}\bar{B}_s]}.\end{aligned}$$

For the particular path y that we chose at (6.12), it was shown by Harris and Git [25] that

$$\sup_{0\le w\le\tau} J(w) = J(\tau)$$

and therefore we have

$$\textbf{sum term} \le e^{tJ(\tau)} \sum_{k=1}^{n_\tau} e^{r\int_0^{S_k} \eta_s^2 - y_s^2\,\mathrm{d}s} e^{[\sqrt{a}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s] + [\frac{1}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}.$$

Proposition 4.2.10 implies that for $0 \le \alpha \le 1$,

$$\tilde{\mathbb{Q}}_t\left(\textbf{sum term}^\alpha\right) \le e^{\alpha t J(\tau)}\,\tilde{\mathbb{Q}}_t\left(\sum_{k=1}^{n_\tau} e^{\alpha r t\int_0^{S_k}[y_s^2 - (\eta_s^t)^2]\,\mathrm{d}s} e^{\alpha[\sqrt{at}\lambda\int_0^{S_k} y_s\,\mathrm{d}\tilde{W}_s] + \alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^{S_k}(\dot{y}_s + \frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\right),$$

and we can transform the sum into an integral by standard techniques (see Kallenberg [26] for example), since the fission times on the spine form a Cox process of rate $2(r\eta_w + \rho)$, as explained in Theorem 6.4.1:

$$= 2e^{\alpha t J(\tau)}\,\tilde{\mathbb{Q}}_t\left(\int_0^\tau e^{\alpha rt\int_0^w [y_s^2-(\eta_s^t)^2]\,\mathrm{d}s} e^{\alpha[\sqrt{\alpha t}\lambda\int_0^w y_s\,\mathrm{d}\tilde{W}_s]+\alpha[\frac{\sqrt{t}}{\sqrt{\theta}}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)\,\mathrm{d}\tilde{B}_s]}\left[r\eta_w^2+\rho\right]\mathrm{d}w\right);$$

Fubini's theorem can be applied to this, and the transformations that worked on the spine term to give (6.27) can here too be applied to arrive at

$$= 2\,e^{\alpha t J(\tau)}\int_0^\tau e^{\alpha^2\int_0^w \frac{a\lambda^2}{2}y_s^2\,\mathrm{d}s} e^{\frac{\alpha^2}{2\theta}\int_0^w(\dot{y}_s+\frac{\theta}{2}y_s)^2\,\mathrm{d}s}\times\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w,$$

$$\leq 2\,e^{\alpha t J(\tau)}e^{\alpha^2 tM}\times\int_0^\tau\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w.$$

We want to take advantage of the fact that the terms in the integral look similar to those already dealt with for the spine term. A first step in this direction is to replace the random factor $rt(\eta_w^t)^2$ at the front of the expectation with the deterministic rty_w^2, and since the value of α will eventually be chosen and fixed the following estimate is sufficient for our purposes.

Lemma 6.6.6 *For all $\alpha > 0$, and for all large enough t,*

$$\int_0^\tau\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w\leq\frac{1}{\alpha}+\int_0^\tau\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w.$$

Proof: Noting that the expectation looks something like $\partial/\partial w\,\tilde{\mathbb{Q}}_t^\alpha\left(e^{-\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$, we shall use integration by parts. From

$$\frac{\partial}{\partial w}\,\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)=\tilde{\mathbb{Q}}_t^\alpha\left(\alpha rt[y_w^2-(\eta_w^t)^2]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right),\tag{6.35}$$

it follows that

$$\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{-\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)=\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$$
$$-\frac{1}{\alpha}\frac{\partial}{\partial w}\,\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right).$$

Integration by parts now proves

$$\int_0^\tau\tilde{\mathbb{Q}}_t^\alpha\left(\left[rt(\eta_w^t)^2+\rho\right]e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w=\int_0^\tau\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w$$
$$+\frac{1}{\alpha}\left[1-\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\right].$$

The exponential decay proved in Theorem 6.6.4 implies $\lim_{t\to\infty}\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)=0$, and this completes the proof □

It follows therefore that for all large enough t,

$$\tilde{\mathbb{Q}}_t(\textbf{sum term}^\alpha)\leq 2\,e^{\alpha tJ(\tau)}e^{\alpha^2 tM}\times\left(\frac{1}{\alpha}+\int_0^\tau\left[rty_w^2+\rho\right]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w\right).$$

We now make some simple over-estimates of the integral. Firstly, it is immediate that

$$\int_0^\tau [rty_w^2+\rho]\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w \le [rt\kappa^2+\rho]\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w$$

since $\left(\sup_{0\le w\le\tau} y_w^2\right)=\kappa^2$. Then, for each $w\in[0,\tau]$, it is true by definition that

$$e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\le e^{\alpha rt(\sup_{0\le w\le\tau}\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s)},$$

and therefore

$$\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\le\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right).$$

Since this holds for all $w\in[0,\tau]$ we can deduce

$$\sup_{0\le w\le\tau}\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\le\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right),$$

which we can use to get:

$$\begin{aligned}\int_0^\tau \tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)\mathrm{d}w &\le \tau\times\sup_{0\le w\le\tau}\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right),\\ &\le\tau\times\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right).\end{aligned}$$

Thus we arrive at a simple upper bound for the sum term: for all $\alpha\in[0,1]$ and all large t,

$$\tilde{\mathbb{Q}}_t(\textbf{sum term}^\alpha)\le 2\,e^{\alpha tJ(\tau)}e^{\alpha^2tM}\left\{\frac{1}{\alpha}+[rt\kappa^2+\rho]\tau\,\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right)\right\}.\quad(6.36)$$

A second application of Varadhan's lemma

We already applied Varadhan's lemma to the term $\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\int_0^\tau[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s}\right)$, and now we show how it can in fact deal with the more complex term $\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right)$ without much more effort.

Once again the observation

$$\tilde{\mathbb{Q}}_t^\alpha\left(e^{2\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right)<\tilde{\mathbb{Q}}_t^\alpha\left(e^{2\alpha rt\tau(\sup_w y_w^2)}\right)$$

shows that the moment condition (6.28) is satisfied and therefore from Varadhan's lemma, Theorem 6.6.3, it follows that:

$$\lim_{t\to\infty}t^{-1}\log\tilde{\mathbb{Q}}_t^\alpha\left(e^{\alpha rt\left(\sup_w\int_0^w[y_s^2-(\eta_s^t)^2]\,\mathrm{d}s\right)}\right)=\sup_{z\in C_0[0,\tau]}\left\{\left(\sup_{0\le w\le\tau}\int_0^w\alpha r[y_s^2-z_s^2]\,\mathrm{d}s\right)-I(z,\tau)\right\}.$$

For any path z, the action functional $I(z,w)$ is non-decreasing in w and therefore

$$\left(\int_0^w\alpha r[y_s^2-z_s^2]\,\mathrm{d}s\right)-I(z,\tau)\le\left(\int_0^w\alpha r[y_s^2-z_s^2]\,\mathrm{d}s\right)-I(z,w),$$

and taking the supremum over $w\in[0,\tau]$ of both sides we deduce:

$$\left(\sup_w\int_0^w\alpha r[y_s^2-z_s^2]\,\mathrm{d}s\right)-I(z,\tau)\le\sup_w\left\{\left(\int_0^w\alpha r[y_s^2-z_s^2]\,\mathrm{d}s\right)-I(z,w)\right\},$$

We now take the supremum of both sides over the set of paths $z \in C_0[0,\tau]$, and interchange the order to obtain:

$$\sup_z \Big\{ \Big(\sup_w \int_0^w \alpha r \big[y_s^2 - z_s^2\big] \, \mathrm{d}s \Big) - I(z,\tau) \Big\} \le \sup_z \sup_w \Big\{ \Big(\int_0^w \alpha r \big[y_s^2 - z_s^2\big] \, \mathrm{d}s \Big) - I(z,w) \Big\} \quad (6.37)$$
$$= \sup_{0 \le w \le \tau} \sup_z \Big\{ \Big(\int_0^w \alpha r \big[y_s^2 - z_s^2\big] \, \mathrm{d}s \Big) - I(z,w) \Big\}.$$

If we compare the term

$$\sup_z \Big\{ \Big(\int_0^w \alpha r \big[y_s^2 - z_s^2\big] \, \mathrm{d}s \Big) - I(z,w) \Big\}$$

with (6.30) from our first application of Varadhan's lemma, it is clear that Euler-Lagrange optimization techniques will result in exactly the same optimal path for this integral, namely $z_s = K_\alpha y_s$ as at (6.32). Furthermore, evaluating the left-hand side of (6.37) shows that we actually have the equality:

$$\sup_z \Big\{ \Big(\int_0^\tau \alpha r \big[y_s^2 - z_s^2\big] \, \mathrm{d}s \Big) - I(z,\tau) \Big\} = \sup_z \Big\{ \Big(\sup_w \int_0^w \alpha r \big[y_s^2 - z_s^2\big] \, \mathrm{d}s \Big) - I(z,\tau) \Big\},$$
$$= \alpha(1 - K_\alpha^2) \Big[\int_0^\tau r y_s^2 \, \mathrm{d}s \Big] - (K_\alpha - (1+\alpha))^2 \Big[\frac{1}{2\theta} \int_0^\tau \big(\dot{y}_s + \frac{\theta}{2} y_s\big)^2 \, \mathrm{d}s \Big],$$
$$< 0 \quad \big(\text{and } = O(\alpha^2) \text{ as } \alpha \to 0\big).$$

Consequently we see that there is no difference in the growth rate between the remaining terms of the spine term and the sum term:

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha \Big(e^{\alpha r t \left(\sup_w \int_0^w [y_s^2 - (\eta_s^t)^2] \, \mathrm{d}s \right)} \Big) = \lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t^\alpha \Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2] \, \mathrm{d}s} \Big) < 0. \quad (6.38)$$

Concluding the upper-bound for $Z_t(\tau)$

We have shown that

$$\tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) \le e^{\alpha t J(\tau)} e^{\alpha^2 t M} \times \tilde{\mathbb{Q}}_t^\alpha \Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2] \, \mathrm{d}s} \Big),$$

and since we clearly have $\tilde{\mathbb{Q}}_t^\alpha \Big(e^{\alpha r t \int_0^\tau [y_s^2 - (\eta_s^t)^2] \, \mathrm{d}s} \Big) \le \tilde{\mathbb{Q}}_t^\alpha \Big(e^{\alpha r t \left(\sup_w \int_0^w [y_s^2 - (\eta_s^t)^2] \, \mathrm{d}s \right)} \Big)$, it follows that:

$$\tilde{\mathbb{Q}}_t\big(Z_t(\tau)^\alpha\big) \le \tilde{\mathbb{Q}}_t(\textbf{spine term}^\alpha) + \tilde{\mathbb{Q}}_t(\textbf{sum term}^\alpha)$$
$$\le e^{\alpha t J(\tau)} e^{\alpha^2 t M} \Big\{ \Big(1 + 2\big[r t \kappa^2 + \rho\big]\tau\Big) \tilde{\mathbb{Q}}_t^\alpha \Big(e^{\alpha r t \left(\sup_w \int_0^w [y_s^2 - (\eta_s^t)^2] \, \mathrm{d}s \right)} \Big) + \frac{2}{\alpha} \Big\}. \quad (6.39)$$

Thus

$$\lim_{t\to\infty} t^{-1} \log \tilde{\mathbb{Q}}_t\big(Z_t(\tau)^\alpha\big) \le \alpha J(\tau) + \alpha^2 M,$$

and the proof of Theorem 6.4.4 is completed. □

Bibliography

[1] Søren Asmussen and Heinrich Hering, *Strong limit theorems for general supercritical branching processes with applications to branching diffusions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36** (1976), no. 3, 195–212. MR 54 #8900

[2] J. D. Biggins, *Martingale convergence in the branching random walk*, J. Appl. Probability **14** (1977), no. 1, 25–37. MR 55 #6592

[3] J.D. Biggins and A.E. Kyprianou, *Measure change in multitype branching*, Adv.Appl.Probab **36** (2004), no. 2.

[4] A. Champneys, S. Harris, J. Toland, J. Warren, and D. Williams, *Algebra, analysis and probability for a coupled system of reaction-diffusion equations*, Philosophical Transactions of the Royal Society of London **350** (1995), 69–112.

[5] B. Chauvin, *Arbres et processus de Bellman-Harris*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 209–232. MR 88a:60151

[6] B. Chauvin and A. Rouault, *Boltzmann-Gibbs weights in the branching random walk*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 41–50. MR 99a:60090

[7] Brigitte Chauvin, *Product martingales and stopping lines for branching Brownian motion*, Ann. Probab. **19** (1991), no. 3, 1195–1205. MR 93b:60188

[8] Brigitte Chauvin and Alain Rouault, *KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees*, Probab. Theory Related Fields **80** (1988), no. 2, 299–314. MR 90b:60113

[9] Brigitte Chauvin, Alain Rouault, and Anton Wakolbinger, *Growing conditioned trees*, Stochastic Process. Appl. **39** (1991), no. 1, 117–130. MR 93d:60138

[10] Joel E Cohen, *Convexity of the dominant eigenvalue of an essentially nonnegative matrix*, Proc. Amer. Math. Soc. **81** (1981), no. 4, 657–658.

[11] E.C.M. Crooks, *Remarks on the Vol'pert theory of travelling-wave solutions for parabolic systems*, University of Bath Preprint.

[12] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Springer, 1998.

[13] Richard Durrett, *Probability: Theory and examples*, 2nd ed., Duxbury Press, 1996.

[14] János Engländer and Andreas E. Kyprianou, *Local extinction versus local exponential growth for spatial branching processes*, Annals of Probability **32** (2004), 78–99.

[15] Yoav Git, *Almost sure path properties of branching diffusion processes*, Séminaire de probabilités, vol. XXXII, Springer, 1998.

[16] Robert Hardy and Simon C. Harris, *Martingales and travelling waves for a finite-type branching diffusion*, University of Bath preprint, January 2004.

[17] ______, *A new formulation of the spine approach*, University of Bath preprint, January 2004.

[18] ______, *A spine proof of a lower-bound for a branching diffusion*, University of Bath preprint, January 2004.

[19] ______, *A spine proof of the BBM large-deviations principle*, University of Bath preprint, January 2004.

[20] ______, *Spine proofs for $\mathcal{L}^p$-convergence of branching-diffusion martingales*, University of Bath preprint, January 2004.

[21] S. C. Harris and D. Williams, *Large deviations and martingales for a typed branching diffusion. I*, Astérisque (1996), no. 236, 133–154, Hommage à P. A. Meyer et J. Neveu. MR 98c:60110

[22] Simon C. Harris, *Branching diffusions*, Ph.D. thesis, University of Bath department of mathematics, 1995.

[23] ______, *Travelling-waves for the FKPP equation via probabilistic arguments*, Proceedings of the Royal Society of Edinburgh **129A** (1999), 503–517.

[24] ______, *Convergence of a 'Gibbs-Boltzmann' random measure for a typed branching diffusion*, Séminaire de Probabilités XXXIV, Lecture Notes in mathematics, vol. 1729, Springer, 2000, pp. 133–154.

[25] Simon C. Harris and Yoav Git, *Large deviations and martingales for a typed branching diffusion, 2*, University of Bath Preprint.

[26] O. Kallenberg, *Foundations of modern probability*, Springer-Verlag, 2002.

[27] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, 1976.

[28] H. Kesten and B.P. Stigum, *Additional limit theorems for indecomposable multidimensional Galton-Watson processes*, Ann. Math. Stat. **37** (1966), 1463–1481.

[29] ———, *A limit theorem for multidimensional Galton-Watson processes*, Ann. Math. Stat. **37** (1966), 1211–1223.

[30] ———, *Limit theorems for decomposable multi-dimensional Galton-Watson processes*, J. Math. Anal. Applic. **17** (1967), 309–338.

[31] A. Kolmogorov, I. Petrovski, and N. Piscounov, *Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem*, Mosc. Univ. Bull. Meth. A **1** (1937), translated and reprinted in P. Pelce, *Dynamics of Curved Fronts* (Academic, San Diego, 1988).

[32] M.A. Krasnosel'skii, *Positive solutions of operator equations*, Noordhoff, 1964.

[33] Erwin Kreyszig, *Introductory functional analysis with applications*, John Wiley and Sons, 1989.

[34] Thomas Kurtz, Russell Lyons, Robin Pemantle, and Yuval Peres, *A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 181–185. MR 98j:60122

[35] Andreas Kyprianou, *Travelling wave solutions to the K-P-P equation: alternatives to Simon Harris's probabilistic analysis*, Ann.Inst.H.Poincare **40** (2004), no. 1, 53–72.

[36] Andreas E. Kyprianou and A. Rahimzadeh Sani, *Martingale convergence and the functional equation in the multi-type branching random walk*, Bernoulli **7** (2001), no. 4, 593–604. MR 2002g:60069

[37] Tzong-Yow Lee, *Some large-deviation theorems for branching diffusions*, Ann. Probab. **20** (1992), no. 3, 1288–1309. MR 93i:60052

[38] Torgny Lindvall, *Lectures on the coupling method*, John Wiley and Sons, 1992.

[39] Quansheng Liu and Alain Rouault, *On two measures defined on the boundary of a branching tree*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 187–201..MR 98j:60123

[40] Russell Lyons, *A simple path to Biggins' martingale convergence for branching random walk*, Classical and modern branching processes (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 84, Springer, New York, 1997, pp. 217–221. MR 1 601 749

[41] Russell Lyons, Robin Pemantle, and Yuval Peres, *Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes*, Ann. Probab. **23** (1995), no. 3, 1125–1138. MR 96m:60194

[42] H.P. McKean, *Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskounov*, Comm. Pure Appl. Math. **28** (1975), 323–331.

[43] J. Neveu, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199–207. MR 88a:60150

[44] J. Neveu, *Multiplicative martingales for spatial branching processes*, Seminar on Stochastic Processes (E. Çinlar, K.L.Chung, and R.K.Getoor, eds.), Birkhäuser, 1987, pp. 223–241.

[45] J. R. Norris, *Markov chains*, CUP, 1997.

[46] Bernt Øksendal, *Stochastic differential equations*, fifth ed., Springer, 2000.

[47] Peter Olofsson, *The $x \log x$ condition for general branching processes*, J. Appl. Probab. **35** (1998), no. 3, 537–544. MR 99k:60214

[48] Zbigniew Palmowski and Tomasz Rolski, *A note on martingale inequalities for fluid models*, Statist. Probab. Lett. **31** (1996), no. 1, 13–21. MR 97j:60177

[49] David Pollard, *A user's guide to measure theoretic probability*, Cambridge University Press, 2002.

[50] L.C.G. Rogers, *Coupling and the tail σ-field of a one-dimensional diffusion*, Stochastic Calculus in Application (J. Norris, ed.), Longman, 1988, pp. 78–88.

[51] L.C.G. Rogers and D. Williams, *Diffusions, Markov processes and martingales*, 2nd ed., vol. 1, Cambridge University Press, 2000.

[52] ———, *Diffusions, Markov processes and martingales*, 2nd ed., vol. 2, Cambridge University Press, 2000.

[53] E. Seneta, *Non-negative matrices and Markov chains*, Springer-Verlag, 1981.

[54] Jan Van Tiel, *Convex analysis: An introductory text*, John Wiley and Sons, 1982.

[55] S. R. S. Varadhan, *Large deviations and applications*, SIAM, 1984.